

# Evaluering – som veiviser?

*Meta-evaluering av forsøksfasen i Den norske kirkes  
trosopplæringsreform*

**Merete Huuse**



Masteroppgave i Didaktikk og Organisasjonslæring

UNIVERSITETET I OSLO

Våren 09

## SAMMENDRAG AV MASTEROPPGAVEN I PEDAGOGIKK

**TITTEL:****Evaluering - som veiviser?**

Meta-evaluering av forsøksfasen i Den norske kirkes trosopplæringsreform.

**AV:**

Merete Huuse

**EKSAMEN:**

Didaktikk og Organisasjonslæring

**SEMESTER:**

Våren 09

**STIKKORD:**

Didaktikk, læringsteori, evaluering.

**SAMMENDRAG:**1) Problemområde/problemstilling:

Formålet med denne oppgaven er å bidra til forskningen på kriteriespørsmålet i evaluering. Min foreløpige hypotese i oppgaven var at evaluering ikke uten videre fører til endring og læring, selv om intensjonen med evalueringer generelt er at de skal vise retning i videre utvikling. Det er interessant å se hvordan evalueringresultater blir benyttet, og hvordan et relativt slett grunnlagsmateriale ofte kan brukes som veiviser. Evalueringsbegrepet kan være sterkt knyttet til begrepet kvalitet, samtidig som materialet evalueringen bygger på ofte undergraver denne forbindelsen. Det har også vært interessant i oppgaveprosessen å se hvordan teoretikere strides om hvorvidt evalueringer bør ha som mål å føre til læring.

Evaluering er nært forbundet med læring. Å se seg tilbake og reflektere over egen praksis er utøvelse av profesjonalitet (Dale 1998). Profesjonalitet i arbeidet, enten det handler om individ eller organisasjon, handler om å legge til rette for refleksjon hvis målet er å utvikle kompetente og lærende organisasjoner. Hvordan læring best kan tilrettelegges i evalueringssammenheng er et lite behandlet tema, som jeg mener å ha gitt mer oppmerksomhet gjennom denne oppgaven.

Jeg har søkt å finne ut av hvordan en evaluering ideelt sett bør utarbeides og gjennomføres for å føre til læring. Mitt organisasjonsteoretiske ståsted har vært en "lærende organisasjon". Problemstillingen i oppgaven er derfor:

**Hvordan bør en evaluering utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i en "lærende organisasjon"?**

Med utgangspunkt i problemstillingen har jeg gjennom evalueringsteori forsøkt å besvare hva som kjennetegner en god evaluering, og hvordan den bør utarbeides for å gi et best mulig grunnlag for læring. Dette har jeg gjort ved hjelp av Evert Vedungs åttepunktprogram for ledelse av evaluering, og hans fasemodell for evalueringer i

offentlig sektor. Jeg har videre belyst betydningen av konsistens og realisme i utviklingen av et program for læring, og vist hvordan dette også har betydning for utviklingen av evalueringer. Til dette har jeg benyttet Erling Lars Dales teori om didaktisk rasjonalitet og hans ulike kompetansenivåer. Jeg har videre latt Lars Qvortrups vitensposisjoner vise hvordan evalueringer bør tilrettelegges for at læring skal finne sted. Med bakgrunn i de nevnte teorier har jeg utviklet en modell som jeg mener viser sammenhengen mellom evaluering og læring.

## 2) Metode:

I denne oppgaven er formålet å finne ut hvordan evalueringer bør utarbeides og gjennomføres for å føre til læring. Jeg har nærmet meg dette på to måter, ved å henvende meg til teori og finne ut hva evalueringsfeltet har kommet til når det gjelder dette, og ved å se på en evaluering i offentlig sektor som nylig er gjennomført. Jeg har på denne måten gått veien om empiri for å se hva som er benyttet av den teorien som finnes på evalueringsfeltet i dag, og hva evalueringsfeltet eventuelt kan lære av den utførte evalueringen.

Evalueringen i reform av trosopplæring er benyttet som en test-evaluering som jeg har belyst evalueringsteori opp mot. Jeg har sett på hvordan evalueringen er utviklet og gjennomført, og hva den mener å ha funnet av kunnskap, for deretter å analysere det jeg har funnet i lys av teori. Dette har jeg gjort ved å analysere det materialet forskerne selv har kommet fram til.

Gjennom analysen har jeg da foretatt en metaanalyse av den evaluering som er foretatt, med spesielt fokus på de funn forskerne selv mener å ha funnet, samt det grunnlag de bygger funnene på. Jeg har besvart problemstillingen ved å analysere evalueringsrapporter fra forskerne ved hjelp av en kritisk hermeneutisk innfallsvinkel i analysen.

## 3) Data/kilder:

Alle evalueringsrapporter som er offentliggjorte på hjemmesiden til reform av trosopplæring på LMSet til Kirken.no er kritisk gjennomgått og sammenlignet, og benyttet

---

som grunnlag for analysen. Gjennom analysen har jeg da foretatt en metaanalyse av den evaluering som er foretatt gjennom hele forsøksfasens fem år. Jeg har hatt spesielt fokus på de funn forskerne selv mener å ha funnet, og det grunnlag de bygger funnene på.

Kildene mine for øvrig er anerkjent evaluerings- og læringsteori. Jeg har søkt å gå så bredt som mulig til verks teoretisk på evalueringsfeltet. Tidsrammen jeg har hatt til rådighet i oppgaven har likevel lagt begrensninger.

#### 4) Resultater/hovedkonklusjoner:

Skal evalueringen ha som mål å føre til læring, eller skal den kun være en beretning eller rapport om hva som er skjedd? Hvor går skillet mellom en god og en dårlig evaluering? Er det noen felles kriterier å forholde seg til?

Vedung har vist at evaluering innebærer i etterkant å oppdage, kartlegge, og bedømme resultater med det formål å oppnå økt selvrefleksjon, dypere forståelse og gjennom dette ta bedre begrunnede beslutninger (Vedung 1998). For å oppnå økt selvrefleksjon og dypere forståelse forutsetter det at det har foregått prosesser av læring. Qvortrup viser med sine vitensposisjoner at refleksjon forutsetter prosesser av læring, som krever ulik type tilrettelegging i organisasjonen, dersom refleksjon og læring skal finne sted.

For at evaluering skal innebære refleksjon over praksis, og føre til dypere forståelse, må evalueringen tilrettelegges for at læring skal finne sted. Det ligger også som en forutsetning i dette, at uten denne tilretteleggingen vil ikke evalueringer fungere etter sin intensjon i henhold til Vedungs forståelse av evaluering. Det er derfor nærliggende å hevde at med bakgrunn i denne forståelsen av evaluering og læring, er det et verdikriterium i all evaluering at de bør ha som mål å føre til læring.

Tilrettelegging for læring er et spørsmål om didaktikk. Godt didaktisk arbeid ved hjelp av god planlegging sikrer kvalitet i utarbeidelse og gjennomføring, og kan gi evalueringen funksjon som verktøy for læring i ”en lærende organisasjon”. For at evalueringer skal fungere etter sin intensjon og bringe aktører og organisasjoner til dypere forståelse og refleksjon over praksis, må den legge til rette for refleksjon på ulike nivåer i ”en lærende organisasjon”. Dette er nødvendig for å kunne gi retning videre.

Jeg har gjennom arbeidet med oppgaven utviklet en modell, som jeg har kalt ”didaktisk evalueringsmodell”. Modellen forener, slik jeg ser det, Vedungs engasjement for de ulike fasene i en evalueringsprosess i offentlig sektor, og mitt viktigste anliggende som jeg beskriver gjennom den didaktiske relasjonsmodellen og Dale og Qvortrups teorier for tilrettelegging av refleksjon og læring; viktigheten av didaktisk arbeid i forberedelse av evalueringsarbeider. En prospektiv holdning og innfallsvinkel til evalueringsfeltet, vil kunne bidra til bedre kvalitet i evalueringsresultater og føre til bedre begrunnede beslutninger i offentlig sektor.

Gjennomgangen av den didaktiske evalueringsmodellen har vist at det skal omfattende tilrettelegging i organisasjonen til dersom evalueringer skal føre til læring. Denne tilretteleggingen av evalueringen er med på å bestemme retningen for evalueringen, hva den skal måle, hva det skal måles med, for og med hvem, og hvordan målingen skal foregå. Arbeidet med de didaktiske kategoriene bidrar derfor til å bestemme de øvrige verdikriteriene i evalueringen.

Testevalueringen kan bare i svært begrenset grad benyttes som belegg for mitt anliggende, siden den ytre validiteten er så svak. Likevel har jeg tillatt meg i analysen å hevde at sterkere fokus på planlegging og tilrettelegging for læring på alle nivåer i organisasjonen, ville betydd at evalueringen i reform av trosopplæring sannsynligvis kunne bidratt til bedre resultat for forsøksfasen, og et bedre utgangspunkt for implementeringen i organisasjonen. Jeg mener derfor at testevalueringen min underbygger mitt anliggende om å vektlegge en mer prospektiv innfallsvinkel til evalueringsfeltet. Det er vanskelig å få en evaluering til

å fungere som virksomhet for læring hvis den ikke er tilrettelagt for det i organisasjonen, og det er vanskelig å få læring til å bli en ønsket endringsprosess hvis de ønskede parametre ikke er definerte i forberedelsesfasen, slik de didaktiske kategoriene kan bidra til.

## Innholdsfortegnelse

<b>1. INNLEDNING.....</b>	<b>10</b>
1.1.2 Hypoteser og problemstilling.....	11
1.1.3 Metode .....	12
<b>2. KUNNSKAP, OPPLEVELSE OG TILHØRIGHET .....</b>	<b>18</b>
2.1 BAKGRUNN FOR REFORMEN .....	18
2.2 EVALUERINGENS MANDAT OG ROLLE.....	21
2.2.1 Evalueringsforskerne .....	23
2.2.2 Valg av metode og kriterier.....	24
2.3 EVALUERINGENS FUNN OG BRUKEN AV DEM .....	26
2.4 OPPSUMMERING.....	31
<b>3. EVALUERING SOM GRUNNLAG FOR LÆRING.....</b>	<b>32</b>
3.1 DEFINISJON AV EVALUERING. ....	32
3.2 HVORDAN UTARBEIDE EN EVALUERING? .....	34
3.2.1 Ledelse av utarbeidelse og prosess.....	34
3.2.2 Evaluering og didaktikk.....	35
3.2.3 Evaluering i "lærende organisasjoner" .....	39
3.2.4 Valg av metode.....	41
3.2.5 Hvem skal evaluere? .....	42
3.3 EVALUERING SOM KUNNSKAPSPRODUKSJON? .....	42
3.3.1 Kriterievalg.....	43
3.3.2 Evaluering som virksomhet for læring .....	44
3.4 HVORDAN GJENNOMFØRE EN EVALUERING? .....	47
<b>4. DEN DIDAKTISKE EVALUERINGSMODELLEN .....</b>	<b>50</b>
4.1 HVORFOR EN DIDAKTISK EVALUERINGSMODELL? .....	51
4.1.1 Forberedelsesfasen.....	53
4.1.2 Gjennomføringsfasen.....	63
4.1.3 Videreutvikling og nykobling.....	64
<b>5. ANALYSE.....</b>	<b>68</b>
5.1 ANVENDELSE AV DEN DIDAKTISKE EVALUERINGSMODELLEN .....	68
5.1.1 Forberedelsesfasen.....	69



---

5.1.2	<i>Gjennomføringsfasen .....</i>	86
5.1.3	<i>Nykobling og metakoblingsprosess.....</i>	91
<b>6.</b>	<b>OPPSUMMERING/KONKLUSJON.....</b>	<b>94</b>
	<b>KILDELISTE .....</b>	<b>100</b>

# 1. Innledning

Evaluering som fenomen strekker seg langt tilbake i tid. Noen har argumentert med at det er et urgammelt påfunn på alder med skapelsen, og at den første dokumenterte evaluering finnes i 1.Mos. 1:31: ”Og Gud så på alt han hadde gjort, og se det var såre godt” (Gray&Jenkins 1990: Vedung 1998). Andre mener det startet i antikken, da forskere og lærde hadde tjeneste ved hoffene og bistod fyrstene med ”råd og dåd” i statsanliggender. Fyrstene tok på denne måten vitenskapen i bruk, som en del av sin virksomhet (Vedung 1998).

Bruken av evaluering som verktøy i kvalitetssikring og utvikling er mye benyttet i vår tid. Det oppfattes som et tegn på kvalitet og sier derfor noe om den virksomheten som utfører dem. Som enkeltpersoner er vi også godt kjent med fenomenet. Vi evaluerer når vi overnatter på hotell, er på kurs, eller møter offentlige og private institusjoner. De fleste av oss deltar i ulike evaluerende virksomheter både i arbeidssituasjonen og i fritiden. Evalueringer er svært forskjellige av størrelse og omfang, og har vidt forskjellig oppbygging og utforming. Begrepet blir benyttet på alt fra tilbakemeldinger på et enkeltstående møte, til evalueringer av store statlige reformer og større forskningsarbeider. I Norge har vi relativt mye erfaring med evaluering som fenomen, og mange av erfaringene fra reformer innen skolesektoren er vel dokumenterte. Både evalueringene i seg selv samt bruken av funnene må vel kunne sies å være utført med vekslende hell.

## *Bakgrunn for prosjektet*

Etter at jeg har hatt ansvar for årsrapporteringer i flere år, og sett hvordan de er blitt behandlet i organisasjonen, har forholdet mellom rapporteringer og evalueringer undret meg. Jeg har sett evalueringer av utviklingsprosjekter, som har vært kun en del av virksomheten, få større oppmerksomhet enn årsrapportene. Dette til tross for at de har hatt dårligere kvalitet i utforming og gjennomføring. Jeg har også opplevd at årsrapportene som var ment å sørge for læring i organisasjonen, på ingen måte har vært tilrettelagt for at læring skulle finne sted. Evalueringer har hatt navn på seg for å være mer enn årsrapporter, mer enn beskrivelser av hva som er utført av virksomheten gjennom et år. De skal fungere som utgangspunkt for veien videre. Det har vært interessant å se hvordan evalueringsresultater har vært benyttet, og hvordan et relativt slett grunnlagsmateriale ofte kan brukes som veiviser i videre

utvikling. Evalueringsbegrepet kan være sterkt knyttet til begrepet kvalitet, samtidig som materialet evalueringen bygger på ofte undergraver denne forbindelsen. Det har også vært interessant i oppgaveprosessen å se hvordan teoretikere strides om hvorvidt evalueringer bør ha som mål å føre til læring.

For meg er evaluering nært forbundet med læring. Å se seg tilbake og reflektere over egen praksis er utøvelse av profesjonalitet (Dale 1998). Profesjonalitet i arbeidet, enten det handler om individ eller organisasjon, handler om å legge til rette for refleksjon hvis målet er å utvikle kompetente og lærende organisasjoner. Hvordan læring best kan tilrettelegges i evalueringssammenheng er et lite behandlet tema, som etter mitt syn burde vies mer oppmerksomhet enn det gjør.

Formålet med denne oppgaven er å gi et bidrag til forskningen på evaluering som fagfelt ved å se nærmere på valg av metoder og kriterier, samt hvordan en evalueringsprosess bør legges til rette for at den skal kunne gi et læringsutbytte. Jeg vil ha særlig fokus på læring/kunnskapsproduksjonen som eventuelt kommer ut av evalueringsarbeidet, og hvilke metoder og kriterier som er valgt, samt hvordan en evalueringsprosess bør legges til rette for å sikre kvalitet og læring.

### **1.1.2 Hypoteser og problemstilling**

Min foreløpige hypotese i denne oppgaven er at evaluering ikke uten videre fører til læring. Jeg er derfor opptatt av å finne ut hvordan en evaluering bør utarbeides og gjennomføres for å sikre at læring skjer. Dette vil jeg gjøre på to måter, ved å henvende meg til teori og finne ut hva evalueringsfeltet har kommet til når det gjelder dette, og ved å se på en evaluering som nylig er gjennomført. Jeg vil på denne måte gå veien om empiri for å se hva som benyttes av den teorien som finnes på evalueringsfeltet i dag, og hva evalueringsfeltet eventuelt kan lære av den utførte evalueringen.

Problemstillingen blir derfor:

**Hvordan bør en evaluering utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”?**

Jeg vil ha en spørrende tilnærming, og hente inn teori der jeg finner det naturlig. Oppgaven vil bli teoretisk, men ved å se på en allerede gjennomført evaluering foretar jeg en meta-

evaluering og gjennom det også en empirisk undersøkelse.

### *Oppgavens oppbygging og innhold*

Oppgavens innhold er organisert om problemstillingen. Den vil styre presentasjonen av temaer og selve innholdet i oppgaven. I kapittel 2 vil jeg redegjøre for evaluering av reform av trosopplæring, hvordan evalueringen i følge forskerne er utviklet og gjennomført, hva forskerne beskriver av funn, og hvilket grunnlag de bygger funnene på. Jeg vil i denne redegjørelsen særlig ha fokus på hvordan evalueringen er tilrettelagt for at organisasjonen skal lære av den, om den bidrar med kunnskapsproduksjon, og hva dette igjen sier om valg av metoder, kriterier og krav til validitet.

Kapittel 3 representerer teoridelen i oppgaven. Med utgangspunkt i problemstillingen vil jeg gjennom anerkjent evalueringsteori i dag forsøke å besvare hva som kjennetegner en god evaluering og hvordan den bør utarbeides for å gi et best mulig grunnlag for læring. Dette vil jeg gjøre ved hjelp av Evert Vedungs åttepunktprogram for evaluering. Jeg vil videre belyse betydningen av konsistens og realisme i utviklingen av en plan for læring, og vise hvordan dette også har betydning for utviklingen av evalueringer. Dette vil jeg gjøre ved hjelp av Erling Lars Dales begrep om didaktisk rasjonalitet og hans ulike kompetansenivåer. Jeg vil videre la Lars Qvortrups vitensposisjoner vise hvordan evalueringer bør tilrettelegges for at læring skal finne sted. Med bakgrunn i de nevnte teorier har jeg utviklet en modell som jeg mener viser sammenhengen mellom evaluering og læring. Modellen bli presentert i kapittel 4, og vil videre benyttes som utgangspunkt for analysen i kapittel 5. Her vil jeg i det hele tatt vurdere om evaluering av reform av trosopplæring er tilrettelagt og gjennomført som en lærende virksomhet på bakgrunn av den teori jeg har valgt å belyse temaet ut i fra. Avslutningsvis, i kapittel 6, sammenfattes det hele i en konklusjon.

Jeg vil ikke i oppgaven gå inn på hva trosopplæring som fenomen er, eller hvorvidt det i det hele tatt er mulig å lære bort å tro, selv om spørsmålet på mange måter opptar meg som pedagog. For meg henger dette nært sammen med forholdet mellom episteme (overbeviser) og doxa (overtale), og det instrumentalistiske mistak (Skjervheim 1968) i pedagogisk sammenheng. Dette ligger imidlertid utenfor denne oppgaven.

### **1.1.3 Metode**

Formålet med all forskning er å gjennomføre en faglig fundert undersøkelse som etterprøver,

bekrefter og/eller fornyer kunnskap (Alvesson 1994). I denne oppgaven er formålet å finne ut hvordan evalueringer bør utarbeides og gjennomføres for å føre til læring. For å finne ut av dette har jeg valgt å benytte evalueringen i reform av trosopplæring som en test-evaluering som jeg vil belyse evalueringsteori opp i mot. Jeg vil se nærmere på hvordan den evalueringen jeg har valgt som forskningsobjekt er utviklet og gjennomført, og hva den mener å ha funnet av kunnskap, for deretter å analysere det jeg finner i lys av teori. Dette vil jeg gjøre ved å se på det materialet forskerne selv har kommet fram til. Holder det vitenskapelig standard? Har de sørget for gode metodiske grep som sikrer validitet i svar?

Alle evalueringsrapporter som er offentliggjort på hjemmesiden til reform av trosopplæring på LMSet til Kirken.no vil bli kritisk gjennomgått og sammenlignet, og benyttet som grunnlag for analysen. Gjennom analysen har jeg da foretatt en metaanalyse av den evaluering som er foretatt gjennom hele forsøksfasens fem år, med spesielt fokus på de funn forskerne selv mener å ha funnet og det grunnlag de bygger funnene på. Jeg vil besvare problemstillingen ved å analysere de evalueringsrapporter som foreligger fra forskerne ved hjelp av en kritisk hermeneutisk innfallsvinkel i analysen.

### *Kritisk hermeneutikk*

Siden denne oppgaven handler om å tolke de tekster som er produsert av evaluerings-forskerne, er det naturlig å knytte seg opp i mot den hermeneutiske tradisjonen i metodevalget.

Hovedtemaet i hermeneutikken er som kjent tolkning av tekst, og det faktum at forståelsen av en del bare kan forstås hvis den sees i sammenheng med helheten, og omvendt. En hermeneutisk tolkning setter et fenomen inn i dets kontekst, det vil si den menneskelige, sosiale, historiske og økonomiske konteksten for å prøve å få tak i selve meningsinnholdet. Et utsagn abstrahert fra sin kontekst kan ikke oppfattes i sin helhet (Kvale 2005). Dette er i samsvar med det holistiske standpunktet som blant annet også gestaltpsykologi og systemteori bygger på; helheten er mer enn summen av delene. I følge Ricoeur fastslår konteksten budskapets mening (Alvesson 1994).

En annen side ved hermeneutikken er forforståelsens betydning i fortolkningsarbeid, noe som skiller den fra fenomenologien. Fenomenologien fokuserer på det deskriptive ved en beskrivelse av en opplevelse og hva som gjør at den oppleves slikt. Mottakeren av beskrivelsen må sette sin for-forståelse i parentes for å få tak i fortellers perspektiv (Kvale

2005). Hermeneutikk ser på beskrivelsens mottaker, med sin for-forståelse, som en 'medkonstruktør' (Patton 2002) eller en 'oversetter'. Ricoeur påpeker den sistnevnte ved å sitere George Steiners utsagn: "Comprendre, c'est traduire" (Alvesson 1994). Dette betyr at fra et hermeneutisk perspektiv, vil ikke bare valg av teori og dens presentasjon i en oppgave være en oversettelse av det som blir lest, men leseren vil i andre omgang også oversette og gjøre teksten til sin egen.

Tema og organisasjonstilknytning i den evalueringen jeg har valgt som objekt kan sies å være belastet av forforståelse og fordom siden det handler om kirke, kirkelige ansatte, troende, eventuelt ikke-troende foreldre og barn. Begrepene vekker bilder i hver forståelseshorisont. Forforståelsen er mangfoldig. Gadamer (1990) snakker om fordom og forforståelse samt om subjektive sider, erfaringer og formeninger. Forståelseshorisonten frembringer forforståelse. Forståelseshorisonten kan forstås som helheten av de bevisste og ubevisste oppfatninger og holdninger en person har. Med meningsutvekslinger (dialog) med det som står utenfor horisonten vil man utvide forståelseshorisonten og dermed også forforståelse og fordommer (Gadamer 1990).

Språk er ved siden av sansing, verktøyet innenfor forståelseshorisonten. Gadamer (1990) tydeliggjør språkets tvetydighet når han henviser til diskreditering av "fordom" som begrep i Hermeneutikken. Selv om fordommen er feilaktig er den en fast bestanddel av vår forståelseshorisont. 'Sannheten' i et hermeneutisk perspektiv er relativ og blir til i dialog mellom forskers sannhet og informantenes sannhet samt fagteoriens sannhet. For å forstå og begripe integriteten av annerledessannheten i en fortolkning må den indre bevisstheten overfor muligheter til misforståelse og feiltolkning skjerpes. I følge Gadamer (1990) er dette hovedinnholdet i den hermeneutiske sirkelen; for-forståelse er en produktiv atferd og ikke reproduktiv.

I følge Alvesson mener Habermas at det ofte ligger en skjult ideologi bak teksten (Alvesson 1994). For å avdekke denne bruker han diskurs som virkemiddel for å undersøke om dialogen er i orden. Diskursen er et hjelpemiddel for å føre kritisk refleksjon med teksten og seg selv som tolker. For Habermas er dette et spørsmål om rasjonalitet. Habermas kritiserer Gadamer for å være for naiv i sitt forhold til dialogen, og får hjelp av Betty som mener at det må skilles mellom tekstens mening; avsenders opprinnelige intensjon, og tekstens signifikans; de tolkninger vi alle har fra vårt ståsted. Den opprinnelige mening og intensjon

med teksten er viktig for Betty, som hevder at man i større grad må balansere mening og signifikans i en tekst tolkning. Avsenders intensjon får vi tak i ved å ærlig underordne oss tekstens opprinnelige poeng i følge Betty (Alvesson 1994).

Metoden som må benyttes for å få tak i meningen i en tekst er en hypotetisk deduktiv metode. Hermeneutikk er en hypotetisk deduktiv metode brukt på meningsfullt materiale (Alvesson 1994). Det settes opp en hypotese for hva teksten vil fortelle, som deretter vurderes og kritisk drøftes for å bekrefte eller avkrefte stemmer. Bak denne metoden ligger forståelsen av at det ikke finnes en sannhet, men flere ulike oppfatninger av hva som er sant.

Vitenskap vil ofte innebære å forholde seg kritisk til gjeldende tolkninger og forklaringer. Uten kritikken av den eksisterende kunnskapen, ville vi neppe se noen videreutvikling av vitenskapene. Derfor er det også en anerkjent virksomhet å drive kritikk av andre forskeres rapporter og publikasjoner (Gunnar Handal 1999). Det er dette jeg gjør i denne meta-evalueringen, selv om siktemålet først og fremst er finne fram til framgangsmåter i evalueringsprosesser som kan sikre at læring skjer. Min viktigste drivkraft er derfor ikke å kritisere forskerne i reform av trosopplæring, men derimot å vise at evalueringer uten læring som mål ikke duger.

### *Meta-evaluering*

Meta-evaluering er som nevnt mitt middel for å svare på problemstillingen. Evalueringsforskning er generelt anvendt forskning som har som formål å fornye og forbedre praksis ved å uttale seg om hvilken verdi et program, et tiltak eller en reform, altså en systematisk handlingsmåte, har (Patton 1990). Metaevaluering har det samme formålet, men målet er ikke rettet mot et program, tiltak eller reform, men heller mot det teoretiske fagfeltet evaluering, eller enkelt- evalueringer som skal iverksettes i fremtiden.

Vedung beskriver metaevaluering som det fjerde laget i sin flersjiktet modell. Meta-evaluering bygger på forestillingen om at selv evalueringsprosesser som er gjennomførte bør evalueres. Dette beskriver Vedung som en etterkontroll. En metaevaluering bør i følge ham deles inn i de samme stadiene som den primære evalueringen gjennom en metaprosess. Også erfaringer fra metaevalueringer skal i følge Vedung tilbakeføres til beslutningstakere for å anvendes. Denne prosessen kaller han ”metakopplingsmekanism”.

Med denne oppgaven vil jeg gjøre et forsøk på å gi et bidrag til denne metakoblingsmekanismen. Jeg regner med at evaluering av forsøksfasen i reform av trosopplæring vil bli fulgt opp med evaluering også av implementeringsprosessen i reformen. Det vil derfor forhåpentlig kunne være nyttig å lese denne oppgaven før evalueringsforskerne eventuelt går videre med arbeidet i implementeringsfasen. I tillegg håper jeg at evalueringsfeltet generelt vil kunne lære noe av evalueringen som er valgt som objekt, slik at evalueringsfeltet styrkes med noe empiri gjennom denne metaevalueringen.

### ***Validitet***

Siden dette handler om en masteroppgave på 30 studiepoeng, har jeg ikke kunnet følge mitt første ønske om å foreta en metasyntese over flere evalueringer. Ved å lande på kun én evaluering sier det seg selv at den ytre validiteten er svært begrenset. I tillegg vil mitt arbeide kun knyttes til evalueringsforskernes egne vurderinger og skriftlige rapporter, med alle de muligheter for feiltolkninger og usikkerhet i funn som det medfører.

Jeg har spesielt i min analyse hatt fokus på metode og kriterievalg, samt i hvilken del av prosessen som forskerne har hatt størst utfordringer i henhold til egne erfaringer med sin forskning. Forskerne har selv vurdert sin egen forskning og gjennom det avslørt både styrke og mangler/svakheter ved sitt arbeid. Jeg har gjennom dette dannet meg et bilde av det forskningsarbeidet som er gjort og benytter dette som utgangspunkt for mine tolkninger. Det er disse tolkningene som blir vurdert i henhold til teori. Jeg har spurt hva forskernes oppsummering av sin egen evaluering egentlig sier om kvaliteten på forskningen og de valg som er gjort når det gjelder utarbeidelse og gjennomføring, og hvordan de har tilrettelagt for å gi et best mulig grunnlag for læring i en lærende organisasjon. Ved ikke å gå dypere inn i det materialet de har hatt som utgangspunkt for sine vurderinger, kan jeg ikke vite at materialet faktisk sier det forskerne hevder at de sier. Dessuten vil jeg kunne feiltolke forskernes rapporteringer uavhengig av om materialet sier det ene eller det andre, ved å lese ut i fra min egen forståelseshorisont. Dette vil selvsagt være utfordrende for den indre validiteten, og vil kunne innebære at jeg framsetter kritikk som ikke er holdbar.

Jeg vil imidlertid så langt det er mulig søke å oppnå best mulig validitet i det jeg presenterer. Dette vil jeg gjøre gjennom å ha en åpen kritisk dialog til min egen forskning, slik den kritiske hermeneutikken legger opp til.



### *Planlegging*

Siden oppgaven er ment å utføres på et halvt år, så startet leseprosessen før jul. Jeg har brukt mye tid på å sette meg inn i de rapportene som forskerne har offentliggjort, og på å snakke med involverte i reformarbeidet for å få et helhetlig bilde av prosjektarbeidet. Jeg er invitert inn i prosjektledelsen og sekretariatets arbeid, og er gitt anledning til å se og vurdere det materialet som forskerne har benyttet hvis ønskelig. Jeg har likevel valgt å konsentrere mitt arbeid om forskernes egne beskrivelser, vurderinger og tolkninger. Dette er valgt på grunn av oppgavens rammer, men også fordi jeg ikke har opplevd det som verken nødvendig eller berikende for min problemstilling å vurdere selv de vurderinger og tolkninger som er gjort. Jeg forutsetter i mitt arbeid at jeg har hatt med seriøse forskere å gjøre, som har opptrådt ansvarlig og pålitelig i forhold til sin rolle. Det er også offentliggjort eksempler på rapporteringsskjemaer som er benyttet og spørsmålsstillinger som er stilt. Derfor mener jeg å ha et rimelig godt grunnlag for å analysere og drøfte forskernes arbeid i forhold til anerkjent evalueringsteori.

## 2. Kunnskap, opplevelse og tilhørighet

Selv om evalueringer er svært forskjellige i utforming og gjennomføring, er det likevel noe som er felles og som har overføringsverdi og relevans for evalueringsfeltet som helhet. Jeg mener at et av felleskriteriene handler om å tilføre læring i en eller annen form. Jeg har derfor valgt å gå inn i en evaluering som har eksplisitt uttrykt i sitt mandat at den er ment å tilføre læring. Siden det også er en opplæringsreform som evalueres har jeg forventning om at den pedagogiske vinklingen er sterkere enn i mange andre evalueringer. Jeg har valgt å se nærmere på hvordan denne evalueringen er utformet og gjennomført når formålet er at det skal komme læring ut av den. I dette kapitlet vil jeg derfor redegjøre for den evaluering jeg har valgt som forskningsobjekt; Evaluering av forsøksfasen i Den norske kirkes Trosopplæringsreform, som er kalt kunnskap, opplevelse og tilhørighet.



Kapitlet vil være orientert om problemstillingen: **Hvordan bør en evaluering utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i en "lærende organisasjon"?**

Jeg vil først si litt om bakgrunn og organisering for reformen generelt for å få et bilde av det materialet som evalueringsforskerne har hatt som utgangspunkt for sitt arbeid, før jeg redegjør for selve evalueringen.

### 2.1 Bakgrunn for reformen

Reformen som evalueres har fått navnet; "Størst av alt", og er omtalt som den største reformen i kirkens historie siden innføringen av konfirmasjonen i 1736 (Helga Haugland Byfuglien 2008). Bakgrunnen for Stortingets iverksettelse av reform av trosopplæring var endringen i opplæringsloven og innføringen av faget KRL i skolen. I et samfunn med et økende kulturelt og religiøst mangfold mente Stortinget at det var viktig at barn og unge fikk opplæring i egen tro. Både for å styrke egen identitet og for å forstå andre kulturer og møte andre tradisjoner med respekt og toleranse (Hauglin m.fl 2008). I tillegg til samfunnets motivasjon for reformen var også mange innenfor kirken misfornøyd med både den

undervisning i kristendom som foregikk i skolen, og den dåpsopplæring som fant sted i kirken. Innenfor kirken opplevde man at oppslutningen om dåpsopplæringen varierte sterkt, flere steder var det få eller ingen tilbud til visse aldersgrupper, og de frivillige barne- og ungdomsorganisasjonene opplevde en klar tilbakegang. I følge evalueringsforskerne kom derfor Stortingets satsing på reform av trosopplæring som et positivt giv inn i kirkelige kretser, og reformledelse og forskere opplevde et positivt engasjement for reformen blant tilsatte og brukere i organisasjonen. Det bør derfor kunne hevdes at reformen og derav også evalueringen i utgangspunktet hadde gode vilkår i organisasjonen som helhet.

### ***Reformens mandat og styringsdokumentenes føringer***

De formelle grunnlagsdokumentene for reformen fulgte ordinær politisk utrednings-, saksbehandlings-, og beslutningsprosess. Reformen skulle i sin helhet være statlig finansiert. Det ble antydnet, men ikke vedtatt i Stortingsbehandlingen, at støtten over en ti års periode ville bli på ca 250 millioner kroner pr år. Spenningen foran hvert statsbudsjett har nok derfor preget forsøksarbeidet. Behovet for nye stillinger skulle avklares gjennom erfaringene fra forsøksarbeidet, og dette skapte naturlig nok stor interesse for reformarbeidet i hele organisasjonen. Det lå også i kortene at det ikke ville bli økninger i andre budsjetttrammer til kirken så lenge denne reformen pågikk, og organisasjonen som helhet ble derfor preget av Stortingets føringer som gikk utenom all annen budsjettmessig struktur. Med penger tildeles også makt, og det må være grunnlag for å hevde at Stortinget gjennom denne reformen også i noen grad forskjøv maktbalansen i organisasjonen. Mye makt ble tildelt reformledelsen.

Hensikten med reformen var fra Stortingets side å sikre alle menighetene i Den norske kirke en mulighet til å tilby en systematisk trosopplæring til alle døpte. Målet er formulert følgende: å utvikle en systematisk trosopplæring som fremmer kristen tro, gir kjennskap til den treenige Gud og gir hjelp til livstolkning og livsmestring for alle døpte i alderen 0-18 år, uansett funksjonsnivå.

Den viktigste føringen for reformen fra Stortingets side var at den skulle være lokalt initiert og forankret. Den skulle gjennomføres over ti år, i første omgang som et femårig prosjektorganisert forsøks- og utviklingsarbeid. Stortinget valgte en ”bottom up”-strategi: få føringer og rammer ble lagt, og menighetene fikk i prinsippet stort spillerom til å eksperimentere med egne trosopplæringstiltak. ” (Hauglin m.fl 2008) Stortinget besluttet

også at forsøksperioden skulle følges av en forskningsbasert evaluering, men la ingen føringer på hvem som skulle ha evalueringsoppdraget eller hvordan det skulle utføres.

### ***Ledelse av reformen***

Organiseringen av forsøkstiden ble derfor i det store og hele av Stortinget overlatt til Kirkerådet, som igjen utnevnte en styringsgruppe som skulle ha prosjektlederansvaret. Styringsgruppen ble bredt sammensatt og opprettet for hele femårsperioden. Oppgavene til styringsgruppen har vært å legge opp en strategisk plan for prosjektperioden, legge de sentrale føringene for alle utlysningsprosesser, vurdere kriterier for søknadsbehandling og vedta tildelingene. Et viktig fokus for deres arbeid har vært å utvikle strukturer rundt menighetenes utviklingsarbeid, som er ment å ha virket kvalitetsfremmende og hvis mål har vært å bidra til å utvikle kirken som en lærende organisasjon (Styringsgruppas rapport 2008). Dette innebærer med andre ord å legge til rette for all rapporteringsvirksomhet som igjen har vært grunnlag for den evaluering som er utført. Dette arbeidet er organisert på 26 møter hvorav seks har gått over to dager. Styringsgruppa har rapportert årlig til Kirkerådet og Kirkemøtet. Kultur- og Kirkedepartementet har mottatt både årsrapporter, muntlige statusrapporter, og fortløpende evalueringsrapporter.

Trosopplæringssekretariatet har forestått den daglige ledelsen av reformen, ut i fra Styringsgruppas vedtak og de føringer som er gitt i tildelingsbrev og politiske vedtak. Det er også de som har sørget for opprettelse og gjennomføring av den forskningsbaserte evalueringen som Stortinget initierte. Det må derfor være lov å si at sekretariatet har hatt stor makt og innflytelse over forsøksarbeidets organisering og gjennomføring, og må derfor ta stor del av ansvaret for ledelsen av arbeidet selv om det formelle ansvaret ligger hos Kirkerådet.

### ***Reformens organisering og omfang***

Forsøksfasen ble organisert på to nivåer, med den sentrale prosjektledelsen øverst og flere lokale styringsgrupper med prosjektledere i hvert av de 156 forsøksprosjektene som reformen tilslutt omfattet. Av 1300 menigheter i Norge var 350 tilslutt involvert i et forsøksprosjekt. Ca. 80% av menighetene søkte om å bli forsøksmenighet i løpet av forsøkstiden. De lokale forsøkene har omfattet mange tusen tiltak og bortimot 300 ansatte og flere tusen frivillige. Det har derfor vært stor involvering av mange i organisasjonen i forsøksfasen. Likevel er det også fremdeles slik at de fleste ansatte og brukere ikke har

deltatt i reformarbeidet i 2009. Det skyldes prosjektorganiseringen som har gått utenom de opparbeidede strukturene i kirken som organisasjon. Dette ble tolket av prosjektledelsen som en føring fra Stortingets ”bottom up”- bestilling, og innebærer at forsøksfasen har foregått ved siden av eller helt utenom den ordinære drift i organisasjonen. Forskerne har omtalt forsøksarbeidet lokalt som ”en stat i staten” der det har vært konflikter, og det kan være grunn til å spørre om ikke hele reformen med sin organisering og ledelse er forstått og opplevd som en ”stat i staten” av tilsatte i kirken som ikke har vært direkte involverte i reformarbeidet. Dette vil jeg komme tilbake til som en faktor jeg mener har innvirkning på organisasjonens evne til å lære av evalueringen og forsøksarbeidet generelt.

Forsøksprosjektene ble etter hvert pålagt en omfattende rapportering som forskerne har benyttet i sin evaluering. Evalueringsforskernes vurdering er at reformen i all hovedsak har vært godt organisert. Samlet ble det bevilget over 375 mill. kroner til gjennomføring av forsøksfasen over Statsbudsjettet, hvorav 288 mill. kroner gikk til lokale forsøk. Det er med andre ord en stor reform som er evaluert, og derav et stort evalueringsoppdrag som er gjennomført.

Jeg har nå sett på bakgrunn og organisering av selve reformen og har derav bedre grunnlag for å vurdere forskernes arbeid med å utvikle, organisere, og gjennomføre evalueringsarbeidet. Jeg vil nå gå nærmere inn på selve evalueringen.

## 2.2 Evalueringsens mandat og rolle

Rammene for evalueringen skaper forutsetningene for evalueringsarbeidet og har derav betydning for utarbeidelse, gjennomføring, og kvalitet i resultat. Det er derfor naturlig å se nærmere på hvilket mandat evalueringen er gitt av oppdragsgiverne og hvilken rolle den har i reformarbeidet.

I forarbeidene til reformen er det understreket at den femårige forsøks- og utviklingsfasen skulle følges av en forskningsbasert evaluering. Evalueringsens mandat og mål er nedfelt i en avtale mellom evalueringsforskerne og Styringsgruppen i reform av trosopplæring av 1. november 2004. I avtalen heter det at evalueringen skal beskrive trosopplæringens utforming, utvikling og konsekvenser og gi informasjon om hvilke erfaringer forsøksmenighetene, de døpte og deres foreldre har med trosopplæringen, samt hvilke

resultater som oppnås. Dette innebærer å belyse konsekvensene av å ha vært forsøksmenighet. Gjennom dette skal evalueringen gi grunnlag for læring for forsøksmenighetene og prosjektledelsen, og for bedre styring av forsøket. Det forutsettes videre i avtalen at forskerne også gjennom evalueringsarbeidet skal gi grunnlag for å vurdere overgangen fra forsøk til varig reform. I tillegg til dette vil prosjektledelsen at forskerne skal gi et bidrag til religionspedagogisk forskning og forskning om trosopplæring (Hauglin m.fl 2008).

Det er derfor ingen liten oppgave som er pålagt forskerne. De har imidlertid i følge prosjektledelsen løst oppgaven tilfredsstillende. Underveis i hele forsøks tiden har evalueringsforskerne gjennom evalueringen lagt fram undersøkelser for å kunne bidra til forbedring av den nasjonale og lokale prosjektstyringen. Dette er gjort gjennom delrapporter og årlige underveisrapporter. Det er også gjennomført casestudier og feltarbeid i syv forsøksprosjekter for å kartlegge og dokumentere hvordan de dømte og deres foreldre opplever reformen. Dette viser en tydelig brukerstyrt evalueringsmetode som jeg vil beskrive nærmere under teoridelen. Rollen forskerne er tildelt er i utgangspunktet tydelig og sterk i forhold til prosjektledelsen. Forskerne er ment å fungere som et verktøy for ledelsen i deres arbeide, og de er i så måte gitt makt og muligheter til å påvirke forsøksarbeidet. Hvorvidt de har hatt denne funksjonen i praksis er et annet spørsmål. Jeg vil se nærmere på dette mot slutten av kapitlet ved å se på hvilke råd som er gitt og hvilke som er fulgt opp av prosjektledelsen.

Evalueringsforskerne har tolket mandatet sitt slik at evalueringen skal fokusere på forsøksfasen og på overgang fra forsøk til reform. Mandatet er sammenfattet i et overordnet spørsmål;

***Har forsøksfasen vært vellykket, og er intensjonen oppfylt?***

Med dette som utgangspunkt har de valgt å undersøke hvilke endringer forsøksfasen har medført for Den norske kirkes trosopplæring. Med endringer har forskerne forstått både kvalitative og kvantitative virkninger som inntreffer, dels som følge av reformens nasjonale virkemidler, dels av lokale forhold, og dels som en følge av andre sider ved samfunnsutviklingen. Forskerne mener å beskrive og forklare hva endringene har bestått i, samt hvordan og hvorfor de skjedde. De har også søkt å forklare reformens innhold og retning ved å plassere den inn i en større teologisk og kirkelig sammenheng.

## 2.2.1 Evalueringsforskerne

Hvem evalueringsforskerne er har betydning for evalueringens utforming og de valg som er tatt vedrørende metode og kriterier. Hvordan de tolker oppgaven og derigjennom utfører oppdraget vil være preget av hvilken kompetanse og erfaring de besitter. Jeg har derfor sett nærmere på hvem forskerne er og hvor mye kompetanse de har på evalueringsfeltet.

Styringsgruppa i reform av trosopplæring ved sitt sekretariat stod som nevnt for etableringen av evalueringsforskere. Evalueringsoppdraget ble utlyst med en åpne prekvalifiseringsrunde. Tre av søkerne ble invitert til en lukket anbudsrunde med forhandlinger. Arbeidsfellesskapet, som består av utdanningsinstitusjonene Diakonhjemmet Høgskole og Det teologiske Menighetsfakultet og ledes av mag.art. Otto Hauglin, fikk oppdraget. Styringsgruppa og sekretariatet beskriver i sin sluttrapport ”Når tro deles”, at arbeidsfellesskapet er forskere med bred erfaring fra evalueringsarbeid og med kompetanse innen samfunnsvitenskap, religionssosiologi og religionspedagogikk.

Forfatterne av den summative sluttrapporten er omtalt med kortfattet CV i boken. De er alle høyt utdannede og har mye kompetanse innenfor det religionsvitenskaplige feltet. Otto Hauglin er den eneste av forfatterne som er omtalt med erfaring fra evalueringsarbeid tidligere. Han har i følge CV og omtale i boken ledet flere større evalueringer på ulike samfunnsområder, herunder i kirkelige sammenhenger. Arbeidsfellesskapet utover forfatterne av sluttrapporten består imidlertid av evalueringsforskere. De har ikke vært med på å forfatte sluttrapporten, men det oppgis i den at evalueringsforskerne har deltatt i evalueringsarbeidet underveis. Alle evalueringsforskerne er i dag knyttet til Diakonhjemmet høgskole. De må derfor sies å tilhøre en institusjon som befinner seg innenfor en kirkelig sirkel. De sitter derfor nært til den kirkelig virkelighet og må derfor hevdes å ha et begrenset utenfra-blikk til evalueringsoppdraget. Forskerne får også veldig god omtale av Styringsgruppa for reformen og deres sekretariat for viktige bidrag i arbeidet; ” Rapportene og møtepunktene med forskerne har vært stimulerende både for den nasjonale ledelse og for utviklingsarbeidet i menighetene” (Styringsgruppas rapport 2008).

Etter å ha undersøkt forskerne nærmere viser det seg at av de som har forfattet sluttrapporten, og også delrapporter og underveisrapporter, har bare en av dem evalueringskompetanse fra før. Denne kompetansen refererer seg også først og fremst til erfaringer med tidligere evalueringsarbeider og ikke faglige kvalifikasjoner. Antakelig for å

bøte på dette har de derfor knyttet seg til et forskningsmiljø som er godt kvalifiserte på evalueringsfeltet. Hvor mye disse forskerne er benyttet i prosessen går imidlertid ikke fram av rapportene. Jeg vil likevel konkludere med at de samlet besitter mye kompetanse på forskning og evaluering, og vil derfor regne med at deres forskning holder god vitenskaplig standard og at vi kan forutsette validitet i deres omgang med data og behandling av materiale. Det er også grunn til å fordre av dem at de kjenner til den evalueringsteori som finnes, og at de valg som er tatt omkring utarbeidelse og gjennomføring av evalueringen er gjort ut i fra erkjennelse av at de sett fra deres synspunkt er de beste for det formålet som er gitt.

### **2.2.2 Valg av metode og kriterier**

Siden vi har slått fast at forskerne har mye kompetanse på evaluering må vi kunne regne med at de har foretatt valget av evalueringsmetode og kriterier ut i fra vitenskaplige vurderinger. De omtaler sin egen forskning og sin innfallsvinkel til forskningsarbeidet veldig kortfattet i den summative evalueringsrapporten. I følge forskerne har deres arbeid vært basert på ulike hermeneutiske og empiriske tilnærminger, herunder både kvalitative og kvantitative metoder. Dette er en svært generell beskrivelse, men vi får nærmere informasjon om valg av metoder hvis vi ser på hva som er utført av forskningsarbeider. Det er gjort dokumentstudier av reformens organisering og gjennomføring, og innholdsanalyser og systematiske gjennomganger av forsøksprosjektenes tiltaks- og årsrapporter. Dette er først og fremst kvalitative studier selv om deler av rapporteringen fra forsøksprosjektene er forsøkt gjort kvantitative. Dette gjelder blant annet materiale som er rapportert om oppslutning om tiltak. Dette materiale vil bli omtalt mer senere i dette kapitlet. Ut over disse analysene og rapporteringene er det gjort studier av reformens støtte struktur og analyser av diverse statistikk. Alt dette materialet er også gjort via rapporteringer og kvantitative tilbakemeldinger. Det er også utført spørreskjemaundersøkelser til de lokale forsøksprosjektene samt et utvalg av ikke-forsøksmenigheter. Dette er det nærmeste vi kommer forsøk på å belyse bieffekter i denne evalueringen. I tillegg til de nevnte undersøkelsene er det også gjort casestudier og en spesialstudie av trosopplæring i samisk sammenheng.

Det er med andre ord gjort et stort arbeid for å dekke ulike aspekter ved forsøkstiden, og mange metoder er tatt i bruk for å skaffe tilveie nødvendig materiale for forskerne. En



forutsetning for evalueringen var at den skulle være både summativ og formativ. Den skulle gi analyser og vurderinger underveis, samt oppsummere og vurdere forsøksfasen. Dette er fulgt opp og gjennomført. Den formative evalueringen er synliggjort gjennom delrapporter og årlige underveisrapporter. Rapportene er fortløpende overrakt Styringsgruppen for trosopplæringsreformen og de er gjort offentlig tilgjengelige. Det er med dette sørget for at det skulle være mulig å lære av resultatene på evalueringen for reformledelsen underveis i hele forsøks tiden. Det at den er offentliggjort på nettet har også gitt andre enn prosjektledelsen mulighet til å sette seg inn i hva evalueringen avdekket, og det er derfor i noen grad gjennom dette lagt til rette for at flere enn prosjektledelsen skal kunne lære av evalueringen underveis i forsøksfasen.

Evalueringsforskerne mener de har hatt grunnlag for å evaluere de ulike prosessene som inngår i forsøks- og utviklingsfasen i reformen med et relativt holdbart datagrunnlag. De hevder selv at i forhold til de aller fleste mål har det metodisk vært mulig å fremskaffe tilstrekkelig grunnlag for å beskrive, og vurdere forsøksfasen (Hauglin m.fl 2008). Dette er imidlertid i følge deres egne slutninger noe hardt tatt i. De beskriver selv relativt store utfordringer knyttet til det datamateriale de har hatt tilgjengelig, som danner grunnlag for vurderinger av forsøks tiden og hva den har gitt av resultater.

Problemene og de metodiske utfordringene mener de i særlig grad har vært knyttet til kvaliteten av vurderingsmaterialet vedrørende status for trosopplæringen ved oppstart av reformen. Det var evalueringsforskernes intensjon å kunne benytte et datamateriale som IKO/NSD fikk i oppdrag å framskaffe i 2003. Dette materiale skulle benyttes som en pretest til å analysere eventuelle endringer. Dette materiale viste seg å bli vanskelig å bruke, fordi registreringsenheten og organiseringen i dette materialet var forskjellig fra det som senere ble brukt i forsøksvirksomheten. Også ulike alderskategorier og annerledes registrering av type tiltak gjorde det vanskelig å benytte materialet. De har likevel benyttet det som en indikasjon på status ved oppstart av reformen, og som en beskrivelse av situasjonen i ikke-forsøksmenighetene i enkelte tilfeller.

I tillegg til utfordringen når det gjelder pretest-materialet er det også funnet svakheter i de rapporteringer som danner grunnlag for evalueringen ellers. Det er blant annet funnet store variasjoner i hva menigheten har lagt til grunn når de har oppgitt antall deltakere for et tiltak, og hvordan de har angitt målgruppe for oppmøte. Disse variasjonene gjør det vanskelig å få

valide tall på deltakelsesprosentene. Også hvordan forsøksmenighetene definerer hva som er et tiltak varierer. Ulike fortolkninger av hva et tiltak egentlig består av, og hvordan det skal rapporteres, gjør det vanskelig å foreta sammenligninger av aktivitetsnivået mellom forsøksprosjektene. Det faktum at det har hersket usikkerhet rundt forståelsen av denne rapporteringen er overraskende med tanke på viktigheten av dette materialet for forskerne. Tilsvarende utfordringer opplever de også å ha når det gjelder måloppnåelse. Prosjektene har vurdert måloppnåelse på en skal fra 1 til 5 hvor 1 er dårligst og 5 er best. Fordi målene er så vidt formulerte og måloppnåelsen så uklart formulert, er det vanskelig å vite hva det betyr når et tiltak skårer 5 i måloppnåelse. Prosjektene har lagt vekt på forskjellige sider i prosjektet i denne vurderingen; f.eks tiltakets gjennomføring, oppslutning, deltakernes opplevelse, virkninger (Hauglin m.fl 2008).

Sett under ett sier evalueringsforskerne selv at disse punktene viser at de må være svært forsiktige med å trekke bastante slutninger som er basert på statistiske opplysninger i tiltaksrapportene. Dette viser til dels store svakheter ved det materialet som har dannet mye av grunnlaget for forskernes vurderinger, og det er derfor alvorlige sider ved forskningsarbeidet som her er avdekket. Er dette mangler som vitner om dårlig forarbeid og svak kvalitetssikring av evalueringsarbeidet fra starten av?

## 2.3 Evalueringens funn og bruken av dem

Etter å ha sett nærmere på hvordan denne evalueringen er utarbeidet og gjennomført, vil jeg nå se på hvilke funn forskerne har gjort underveis i forsøks tiden og hvordan disse funnene er behandlet av prosjektledelsen. Dette kan vise om evalueringen har lyktes i å skape muligheter for læring i prosjektorganisasjonen og organisasjonen som helhet. På den måten kan vi analysere evalueringens utarbeidelse og tilrettelegging med tanke på om den faktisk har ført til læring. Jeg har sett på hva som er avdekket i underveis rapportene, og videre hvordan det igjen er fulgt opp i de følgende rapporteringer og i den summative rapporten.

Gjennom hele forsøks tiden har forskerne avdekket svakheter ved reformarbeidet som er rapportert om. Flere av svakhetene er forbedret av prosjektledelsen. Det viser nærmere analyser av rapportene. Men rapportene viser også at det er deler av rapporteringene som går igjen i alle underveis- og delrapporter gjennom hele forsøksperioden. Dette sier noe om hvilken plass evalueringen har hatt i reformarbeidet, og slik jeg ser det også i hvilken grad

forsøksarbeidet har hatt muligheter for å lære av den. Jeg har sett nærmere på hvilke endringer som er rapportert flere ganger, fordi jeg mener det er et tegn på om følgeforskningen har lyktes eller ikke.

### ***Rapporterte funn som er fulgt opp***

Flere av funnene forskerne har gjort har hatt betydning for utviklingen av hele forsøksfasen. Jeg vil i det følgende oppsummere hvilke funn som er rapporterte og faktisk fulgt opp av reformledelsen. Disse viser etter mitt syn at følgeforskningen har hatt en funksjon.

### **Tildeling av forsøksmidler**

Styringsgruppa og prosjektsekretariatet har i følge forskerne lagt rammene og gjennom ulike styringsimpulser påvirket utformingen av reformen. En viktig del av iverksettingen var å fastlegge kriterier for tildelingen av forsøksmidlene og vurderinger av innkomne søknader. Forskerne påpekte allerede i delrapport 1 at det var ønskelig med større grad av innsyn, skriftlighet og begrunnelser i søknadsbehandlingen. Fraværet av eksplisitte kriterier for å prioritere søknader foran andre, gjorde det vanskelig å etterprøve tildelingene. Dette er et område som prosjektledelsen har tatt tak i og forsøkt å utbedre gjennom alle søknadsrundene. Her har forskerne hatt en tydelig stemme inn i arbeidet, og anbefalingene deres er fulgt. De konkluderer i den summative rapporten med at korrigeringer underveis gjorde senere tildelinger mer målrettet.

### **Bemanning i sekretariatet**

Forskerne meldte også behovet for å styrke det sentrale sekretariatet. Dette ble av forskerne blant annet grunnlagt i det planlagte kompetansenettverket, som til da ikke var satt gang. Forskerne har i alle rapporteringene balansert sin rolle med å påpeke utfordringer og mangler, samtidig som de har uttrykt medfølelse overfor et overarbeidet sekretariat. Til dels sterk kritikk av reformarbeidet er neddempet i lys av dette. De har lyktes i å påvirke reformledelsen til å prioritere mer midler til det sentrale sekretariatet. Sekretariatet er gjennom forsøksfasen styrket med flere stillinger.

### **Utviklingsprosjekter og ulike støttefunksjoner**

Det ble etablert en rekke støttefunksjoner for forsøksvirksomheten. Etablering av en veiledningstjeneste kalt mentorordningen er en av dem. Ordningen representerer en ny type veiledning i Den norske kirke (Hauglin m.fl 2008). Hvert prosjekt har vært tildelt en mentor

som har hatt i oppgave å følge prosjektet gjennom forsøksperioden, og gi råd og veiledning til prosjektledelse og prosjektmenighet. Ordningen er evaluert og fulgt nært av forskerne i hele forsøksperioden, og anbefalinger som er gitt underveis i utvikling av ordningen er i stor grad fulgt. Forskerne mener det er grunnlag for å si at man har lykket med å rekruttere et kompetent og motivert mentorkorps, og at mentorenes innsats etter alt å dømme har bidratt konstruktivt til kvalitetssikring av det lokale prosjektarbeidet.

Opprettelse av kompetansenettverk var et viktig insitament for reformen, og var også direkte bestilt av Stortinget. Hensikten med en slik etablering var å kople sammen kompetanse i ulike kirkelige og samfunnsvitenskaplige miljøer, for å heve kompetanse og kapasitet på kirkelig opplæring. Nettverket trådte i funksjon i 2005, og forskerne beskriver dets utvikling som et relativt løst strukturert nettverk. Det har manglet en tydelig strategi for arbeidet med samordning og utvikling av kompetanse, og har etter forskernes oppfatning bare i begrenset grad fungert som støttestruktur for gjennomføring og kvalitetssikring av reformen. Det må likevel omtales under rapporterte funn hos forskerne som er fulgt opp, siden det ble etterspurt av forskerne i starten av forsøksfasen og forsøkt fulgt opp av reformledelsen i ettertid, selv om de i følge forskerne ikke helt har lykket.

Også nasjonale og regionale utviklingsprosjekter er omtalt og etterspurt tilsvarende av forskerne med samme følger. Utviklingsprosjektene har ivaretatt mange av Stortingets målsettinger om å involvere nye aktører i reformen, bidra til trosopplæring for barn med særskilte behov, og utvikling og utprøving av metoder og materiell. Sentrale aktører her har vært kristne barne- og ungdomsorganisasjoner, fagmiljøer i høyskolesektoren og kirkelige utdanningsinstitusjoner. Forskerne mener at det er lite grunn til å betvile at det er satt i gang et viktig forsknings- og utviklingsarbeide gjennom dette som de understreker må anvendes og gjøres tilgjengelige inn i det videre reformarbeidet. De argumenterer derfor for sterkere koplinger mellom utdanningsinstitusjonene og forsøksvirksomheten for å systematisere og bearbeide erfaringer i prosjektene.

Forskernes råd er fulgt ved etableringen av disse støttestrukturene, selv om forskerne i sluttrapporten påpeker flere utfordringer knyttet til gjennomføringen av dem. Generelt kan disse utfordringene tolkes som et uttrykk for at det er satt i gang og utviklet mange gode tiltak som skulle fungere som støttestrukturer, men som i mangel av en helhetlig organisering og involvering fra prosjektledelsens side skaper utfordringene som forskerne

har avdekket. Kunne disse utfordringene vært unngått ved bedre planlegging både i forhold til forsøksarbeidet og utarbeidelse av evalueringen?

### ***Rapporterte funn som ikke er fulgt opp***

Vi har nå sett at forskerne har hatt påvirkning gjennom sin følgeforskning. Flere av de rapporterte og avdekkede svakheten ved forsøksfasen er bearbeidet på grunn av forskerne. I det følgende oppsummeres de funn som er rapportert flere ganger gjennom forsøksfasen, men som likevel ikke er tatt til følge.

### **Tiltak og oppslutning**

I følge forskerne er det gjennomført et mangfold av aktiviteter som har lagt grunnlag for lokal utprøving, erfaringsinnhenting og læring i forsøksfasen. Men spørsmålet om forsøksfasen har vært vellykket ut i fra et oppslutningsperspektiv er ikke enkelt å besvare. Dette er knyttet til svakheten i rapporteringsmaterialet som er nevnt tidligere, men også det faktum at det ikke forelå uttalte målsettinger for omfang og tall på aktiviteter en ønsket å oppnå når forsøksfasen startet. Det manglet derfor kriterier å vurdere resultatene opp i mot. Dette ble avdekket tidlig av forskerne men er ikke gjort noe med av ledelsen. Kan det skyldes at det ikke har vært hensiktsmessig å skape denne type kriterier i etterkant, når det ikke var på plass ved starten av forsøkstiden? Kan det være et tegn på at det har vært benyttet for lite tid på planleggingsfasen i reformarbeidet, og at forskerne har kommet for sent inn i prosessen til å påvirke utformingen av rapporteringene? Svakhetene i dette har naturlig nok fulgt hele forsøksarbeidet, og har hatt konsekvenser for både reformarbeidet og evalueringen av det.

### **Trosopplæringens innhold**

Det er videre avdekket at reformen mangler en nærmere bestemmelse av hva som skal være trosopplæringens innhold. Det har foregått en lokal teologisk refleksjon om dåp og opplæring og det er utviklet en ny tenkning om læring i mer sosiokulturell retning der læring skjer i et praksisfellesskap, og ikke minst er synet på barn og unge som subjekter, deltakere og sentrale aktører en vesentlig endring i følge forskerne. Men selve innholdet i trosopplæring er likevel ikke bestemt gjennom forsøksarbeidet. Siden innholdsbestemmelsen skal danne grunnlaget for en systematisk trosopplæring, er dette en stor utfordring for prosjektledelsen i forhold til å evne å innfri målet for reformen. Forskerne hevder i sin rapport at erfaringsmaterialet så langt ikke kan si noe om hvordan en systematisk

trosopplæring for hele aldersgruppen bør se ut. Hva sier dette funnet i realiteten om hele forsøksfasen?

Forskerne konstaterer at forsøksmenighetene langt på vei har videreført, forankret og videreutviklet menighetens trosopplæring med grunnlag i egne, lokale tradisjoner for tros- og dåpsopplæring, og slår derfor fast at vi har å gjøre med en kontekstuell reform. Dette kan vel sannsynligvis også forklare hvorfor det er en utfordring å komme fram til en felles innholdsbestemmelse. Uten felles føringer, og uten at det parallelt med de lokale forsøkene ble gjort arbeid for å sikre en helhetlig tenkning, kan det vel heller ikke forventes. Den felles innholdsbestemmelse som det her siktes til er i sum den læring som skulle komme ut av forsøks tiden, og som reformledelsen har hatt som målsetting å nå. Det er også dette målet forskerne skal beskrive gjennom sin evaluering om forsøksfasen har nådd.

### **Uklart skille mellom reform og forsøk**

Det viktigste funnet som representerer utfordringer for reformen er likevel etter forskernes oppfatning at prosjektledelsen har valgt å gjennomføre en forsøksfase *samtidig* som man har startet reformimplementeringen. Med tanke på at forsøksfasen nettopp er ment å innebære prøving og feiling i forkant av innføring av reformen som helhet er det et viktig anliggende. Evalueringsforskerne mener prosjektledelsens håndtering av dette har svekket forsøksdelen av reformen. De har også påpekt svakheten ved dette fra første til siste rapport, så ledelsen har hatt rikelig anledning til å endre praksis underveis. Det faktum at alle stillinger som er besatt i forsøks tiden er gjort til faste stillinger har ikke gjort dette enklere. Det har skapt klare begrensinger i forsøksperioden og lagt tydelige føringer for reformens implementering.

Forskerne hevder at styringsgruppa i for liten grad har instruert prosjektsekretariatet til å drive mer planmessig forsøksvirksomhet. Det har, etter forskernes mening, vært lagt altfor stor vekt på en stadig økende implementering. De er kritiske til at man ikke har utviklet og gjennomført adekvate tiltak for å aggregere forsøks erfaringene til anvendt kunnskap for det videre reformarbeidet. De er blant annet av den grunn opptatt av at det videre arbeidet med plan for trosopplæring i Den norske kirke ikke må forseres. Forskerne har på dette området hatt en tydelig stemme og gitt svært kritiske tilbakemeldinger til prosjektledelse gjennom hele forsøks tiden. Likevel har de ikke nådd fram med synspunktene sine. Dette vitner om en noe svak posisjon i forhold til prosjektledelsen, og kan si noe om hvilken rolle evalueringsforskerne har hatt i forsøksarbeidet. Om denne svikten i myndighet skyldes ledelsen eller forskerne selv vil jeg gå nærmere inn på i analysen. Det er imidlertid min

---

foreløpige hypotese at kvaliteten i evalueringens utarbeidelse har betydning for hvilken rolle den senere får i organisasjonen.

## 2.4 Oppsummering

Denne redegjørelsen av ”Evaluering av forsøksfasen i Den norske kirkes Trosopplæringsreform” har vært orientert rundt hvordan denne evalueringens utarbeidelse og gjennomføring er utført, og hvordan den er tilrettelagt for å føre til læring i organisasjonen. Det er innledningsvis gitt et bilde av den reformen som evalueres for bedre å kunne vurdere forskernes arbeid. Deretter beskrives evalueringens rammer og vilkår, samt hvordan forskerne omtaler sitt arbeid i forhold til utarbeidelse og gjennomføring. Som avslutning på kapitlet beskrives funn og hvordan de er behandlet i organisasjonen og av ledelsen for reformen. Dette mener jeg forteller noe om hvilken rolle evalueringen og forskerne har hatt i forsøksfasen, som etter mitt syn igjen sier noe om kvaliteten på selve evalueringen.

Det er til dels alvorlig kritikk forskerne retter mot utarbeidelse og gjennomføring av forsøksfasen. De sier blant annet at de er usikre på om det er grunnlag for å si at forsøksfasen har vært vellykket. Dette fordi de hevder at den ikke er godt nok utnyttet til testing av ulike tiltak, og at de forsøkene som er gjennomførte ikke er godt nok systematiserte med tanke på å hente læring ut av dem. Forskerne legger ansvaret for dette på prosjektledelsen som etter deres syn har startet implementeringsprosessen *samtidig* med forsøksfasen. Dette mener de har svekket forsøkstiden og hindret fokus på utprøving og systematiseringsarbeide som kunne bidratt til å innholdsbestemme og utvikle trosopplæringen. Forskerne sier også selv at de må være svært forsiktige med å trekke bastante slutninger som er basert på statistiske opplysninger i tiltaksrapportene.

Utfordringene som her avdekkes mener jeg sier noe om forberedelsesfasen i reformen og forberedelsen av evalueringen. Jeg mener det er grunn til å spørre om ikke disse utfordringene kunne vært avdekket og gjort grep for å hindre gjennom grundigere forarbeide i utarbeidelsen av evalueringens rapporteringsgrunnlag. Hvordan forberedelsen kunne vært gjort bedre vil jeg søke å finne svar på hos evalueringsteoretikere. I det neste kapitlet vil jeg derfor se nærmere på hva teori kan fortelle om hvordan en evaluering best bør utarbeides for å legge til rette for læring, og hva det igjen innebærer for en evaluering som denne.

### 3. Evaluering som grunnlag for læring

I kapittel 2 har jeg redegjort for evaluering av trosopplæring i den norske kirke som jeg har valgt som min test-evaluering i denne oppgaven, og som jeg vil gjennomføre en meta-evaluering på. Arbeidet med evalueringen har reist flere spørsmål som jeg i dette kapitlet vil søke å finne svar på i relevant evaluerings- og læringsteori. Utvalget av teori vil være orientert om problemstillingen: **Hvordan bør en evaluering utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”?** Jeg vil starte med å se på hva teoretikere sier om utarbeidelse og gjennomføring av en evaluering, og hva det innebærer for ”lærende organisasjoner”. Derest vil jeg se på hva tilrettelegging av læring innebærer for evalueringer med tanke på at de skal ha som funksjon å føre til læring. Men først vil jeg avklare hva jeg legger i begrepet evaluering.

#### 3.1 Definisjon av evaluering.

”Detta har berett mig det allra största besvär och bereder mig fortfarande det allra största besvär: att inse att det är av oändligt mycket större betydelse vad tingen heter än vad de är” (Friedrich Nietzsche, Den glada vetenskapen: Vedung 1998: s 19).

Kritikere hevder at evalueringenes viktigste betydning er begrepet, og det faktum at det kan sies at evalueringen er foretatt i omgivelsene rundt. Evalueringsbegrepet er omfattet av mye prestisje og status, og det blir ofte brukt som et mål på kvalitet at de i det hele tatt gjennomføres, selv om grunnlaget i innhold og gjennomføring kan være heller tynt. I følge Peder Haug er tendensen at enhver evaluator har sin egen private definisjon av begrepet, og at det derfor finnes et uttall forskjellige definisjoner og betydninger av ordet (Haug og Schwandt 2003). Forskerne i reform av trosopplæring oppgir ikke sin definisjon av begrepet i sine rapporter. Det kan være mange grunner til det, blant annet at det kan være lite kommuniserbart med definisjoner i rapporter som har som mål å skulle leses av et bredt spekter av mennesker. Likevel kan det være viktig for leserens forståelseshorisont å være kjent med hva forskeren legger i begrepet.

Jeg har valgt å støtte meg til Evert Vedungs definisjon av evaluering, og begrepet vil i det neste bli forstått som; ”noggrann efterhandsbedömning av utfall, slutprestasjoner eller



förvaltning i offentlig verksamhet, vilken avses spela en roll i praktiska beslutssituationer” (Vedung 1998).

”Att nogrann bedöma” innebär i følge Vedung å fastsette verdien av det du bedømmer. Evaluering defineres derfor som en aktivitet som systematisk skiller det verdifulle fra det verdiløse (Schriven 1991:1). Evaluering er derfor tilbakeskuende på det som har skjedd, og ikke i følge Vedung ”prognosmakeri och antebedomning”. Evaluering ser derfor tilbake og bedømmer det som er eller har vært for å gi veiledning framover. Evaluering innebærer derfor i etterkant å oppdage, kartlegge, og bedømme resultater med det formål å oppnå økt selvrefleksjon, dypere forståelse og gjennom dette ta bedre begrunnede beslutninger (Vedung 1998). I dette ligger det, slik jeg ser det, at evaluering innebærer læring. For å oppnå dypere forståelse og økt selvrefleksjon har det foregått prosesser av læring, som igjen krever tilrettelegging for at de skal finne sted. Hvordan denne tilretteleggingen bør gjøres og hvordan du best sikrer at denne læring skjer, er det jeg søker å finne svar på i denne oppgaven.

Tilretteleggingen av læring må ta hensyn til at evalueringer er ulike i form og innhold. Det er derfor nødvendig å kjenne til at evalueringer kan være både formative og summative (Patton 1990: Østerud 2006). Formative evalueringer gjennomføres for å forbedre en læringsprosess, og den er ment å forme eller eventuelt omforme læringsprosessen. Denne formen for evaluering løper parallelt med læringsprosessen og må derfor betraktes som en del av den. Formative evalueringer er derfor følgeforskning slik som evalueringen av reform av trosopplæring, og siden målet med denne type evaluering er å forme, forbedre, evt. omforme læringsprosesser er det også mitt fokus i oppgaven hva slags læring som er kommet ut av evalueringsprosessen. Formative evalueringer er mest opptatt av å se helhet i en bestemt kontekst og støtter seg derfor i hovedsak til kvalitative data, men kan også eventuelt i tillegg begrunnes i kvantitative data.

Summative evalueringer retter seg mot produktet av læringsprosessene og tar sikte på å undersøke om det materiale som foreligger gir bedre læringsutbytte enn noe annet. De summative evalueringene er opptatt av kontrollerte sammenlikninger, generalisering og relativt store utvalg, og må derfor være forankret i kvantitative data. Siden målet med summative evalueringer er sammenlikninger og generaliseringer, stilles det store krav til validitet i materialet. Deler av evaluering av reform av trosopplæring er summative, og skal derfor vurderes som det (Østerud 2006). Læring i organisasjoner lar seg studere både

gjennom en summativ og en formativ evaluering, og dette er også forsøkt gjort i evaluering av reform av trosopplæring i DNK.

## 3.2 Hvordan utarbeide en evaluering?

Erfaring viser at evalueringer generelt er svært forskjellige i form og innhold. Hvordan en evaluering skal utarbeides, og hva som er den beste framgangsmåten i det enkelte tilfelle er det i følge forskere ikke enkle svar på. Det foreligger imidlertid ulike modeller og redskaper som kan bidra i arbeidet med å utvikle en evaluering, og som kan bidra til å sørge for best mulig kvalitet i evalueringens resultat. Jeg vil i det følgende beskrive hvordan en som leder bør gå fram i utviklingen av en evaluering, og hvilke forhold som bør ha spesiell oppmerksomhet i denne prosessen. Jeg vil særlig dvele ved hvordan evalueringer bør tilrettelegges i organisasjonen gjennom en formulert plan for læring, og hva dette sier om utarbeidelse og gjennomføring. Jeg vil også se på hva en god utarbeidelse vil si for hvilket organisasjonsteoretisk ståsted du har, hvilken metode du velger, og hvem du bør velge til å evaluere.

### 3.2.1 Ledelse av utarbeidelse og prosess

Vedung har utarbeidet et åttepunktprogram for hvordan man som leder kan nærme seg en evalueringsprosess. Han mener det hele starter med det overgripende motivet bak evalueringen, det han kaller ”syftesfrågan”. I offentlig sektor er evalueringer, slik vi har sett i reform av trosopplæring, ofte bestillingsverk fra f.eks politiske myndigheter. I disse tilfellene er behovet for å finne fram til det egentlige motivet bak bestillingen viktig for å kunne svare på bestillingen. Spørsmål som hvorfor bestilleren vil ha saken utredet og hva den skal benyttes til blir da vesentlig å få svar på. Dette bør så ha betydning for utarbeidelsen av evalueringen, valg av evalueringsmetode, organisering, og ikke minst hvem som skal utføre evalueringen. Hvordan bestillingen av evalueringen er oppfattet vil få følger for det videre arbeidet med utformingen av evalueringen. Derfor er starten av en evalueringsprosess spesielt viktig, og blir derfor et viktig anliggende i analysen.

De øvrige punktene i åttepunktprogrammet til Vedung gjelder indre, innholdsmessige perspektiver. De spør bl.a. om det er sammenheng mellom organisering av evalueringen, intensjonene til virksomheten, og det evalueringen sikter mot. Hvilke kriterier skal benyttes,

og hvilke faktorer skal resultatet forklares med? Dette viktige grunnlagsarbeidet beskriver Vedung som indre, innholdsmessige perspektiver som har avgjørende betydning for kvaliteten på evalueringens produkt. Det er denne delen av evalueringsprosessen som evaluering av reform av trosopplæring etter mitt syn burde lagt mer vekt på, og som jeg vil komme tilbake til i analysen. Disse perspektivene er nært knyttet opp i mot de didaktiske kategoriene som vil bli presentert i punkt 3.2.2.

De indre anliggende av organiseringen rundt en evaluering er i følge Vedung også avhengig av hvordan organisasjonen som helhet forholder seg til den. Evaluering av en virksomhet eller deler av den, kan ikke isoleres rundt selve evalueringsobjektet. Kvaliteten på evalueringen vil være avhengig av en nærhet til organisasjonens øvrige arbeid og organisasjonens øvrige medarbeidere. Dette vil jeg beskrive nærmere i det følgende.

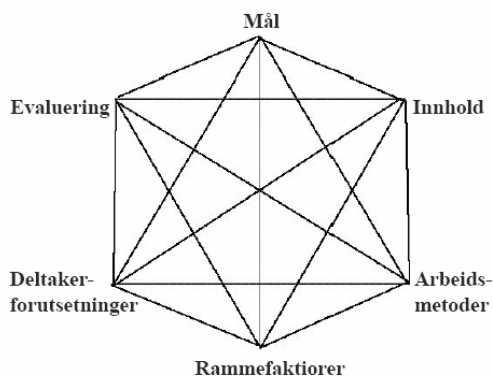
### **3.2.2 Evaluering og didaktikk**

Hvordan en evaluering skal tilrettelegges for å føre til læring i en organisasjon er et spørsmål knyttet til didaktikk. Med didaktikk menes teori om hva som skal læres, om hvordan innholdet skal organiseres og læres, og svar på hvorfor, dvs. legitimere mål, innhold, metode og vurdering (Dale 1998).

Som utgangspunkt for didaktisk arbeid er det naturlig å forutsette at det foreligger en plan for den læring som skal finne sted. Dette er et nødvendig redskap for å systematisere og organisere opplæring. I Dales definisjon av didaktikk forstår jeg vurdering som likestilt med evaluering. Det sees som nødvendig forutsetning for et godt didaktisk arbeid at evaluering foretas, og gis naturlig plass i en opplæringsplan. Like fullt bør det etter mitt syn også utvikles en plan for hvordan opplæringen skal vurderes/evalueres. Dette vil bidra til å sikre kvalitet og ikke minst evalueringens plass i organisasjonen.

Hvordan en evaluering er organisert i en virksomhet, og hvilken plass den har i organisasjonens øvrige arbeid, har betydning. Hvordan dette er organisert har følger for om evalueringen evner å føre til læring. Flere har påpekt betydningen av dette, men jeg velger å benytte den didaktiske relasjonsmodellen til å belyse sammenhengen. Det er naturlig å se evaluering som en del av en prosess for læring og utvikling, som kan synliggjøres via denne modellen. Den didaktiske relasjonsmodellen tar utgangspunkt i kategoriene mål, innhold,

arbeidsmetoder, rammefaktorer, deltakerforutsetninger og evaluering (Bjørndal og Lieberg, 1978). Se figur.



Hensikten med en slik modell har vært å angi hva slags faktorer og relasjoner som bør analyseres i forbindelse med læringstiltak. Hovedbudskapet i modellen er at læring sikres best når det er god sammenheng og indre konsistens mellom de seks kategoriene. Evaluering er en naturlig del av læring generelt, men den må inngå i en helhet

dersom målet er å sikre et læringsutbytte. Ledelse av en evaluering innebærer å ta hensyn til disse faktorene. Det er sannsynligvis lite læringseffekt av en evaluering som blir iverksatt av ledelse, gjennomført litt ”utenfor” den daglige drift, og siden presentert som noe virksomheten skal lære av. Jeg mener at den didaktiske relasjonsmodellen viser at for at evalueringer skal føre til læring i en organisasjon er det nødvendige å sørge for at den er en tydelig del av organisasjonens øvrige arbeide. Dette innebærer også at behovet for evaluering, prosessene rundt utarbeidelse, gjennomføring og behandling av utfall, bør implementeres i organisasjonen som helhet før den iverksettes. Det som her nevnes mener jeg har relevans for evalueringens rolle i organisasjonen og hvordan involvering og tilrettelegging av evalueringen bør skje med tanke på senere implementering. Jeg derfor komme tilbake til dette i analysen.

For å sikre dette didaktiske og helhetlige perspektivet, og at evalueringen faktisk fører til læring, så er som nevnt en formulert plan for hvordan evalueringen skal utføres og hvilken rolle den skal ha i virksomheten nødvendig. Jeg vil videre se på hvordan denne planen bør utarbeides.

### ***Didaktisk rasjonalitet***

Didaktisk rasjonalitet er et viktig begrep i sammenheng med utviklingen av en evaluering. Didaktisk rasjonalitet betyr en evne til å vurdere konsistensen i en formulert plan for læring (Dale 1998). I følge Dale knytter begrepet seg til tre grunnleggende temaer i kombinasjon med tre forskjellige kompetansenivåer. For det første er didaktisk rasjonalitet i en utviklet plan for opplæring, enten den er konstruert for undervisning eller for en organisasjons

selvlæring, et spørsmål om konsistens. Det innebærer, slik også den didaktiske relasjonsmodellen viser, at det skal være en sammenhengende relasjon mellom planens mål, innhold og metodevalg. I evaluering av reform av trosopplæring er målet uttrykt i flere delmål, og samlet i et felles overordnet spørsmål; *har forsøksfasen vært vellykket?* Hvis mål, innhold og metodevalg skal være i overensstemmelse i dette, betyr det at evalueringens tilrettelegging/forberedelse for innhenting av data har hatt som mål å besvare dette spørsmålet best mulig. Det forhold at forskerne har valgt et annet mål enn det som er lagt for reformen, reiser også spørsmålet om konsistens i henhold til eksterne og interne aktører i organisasjonen. I følge Dale blir gruppen av medlemmer i en organisasjon en organisert enhet om medlemmene utfører de ulike funksjoner de er satt til å forvalte i lys av felles mål. Kollegaene generaliserer en holdning til organisasjonene som enheten i den sosiale prosessen de deltar i. Med referanse til Mead (1934) benytter Dale begrepet ”den generaliserte andre” på holdningen til det kollegiale fellesskapet som helhet (Dale 1998). For at evalueringen skal bli en del av ”den generaliserte andre”, og generaliseringene som kollegene gjør, betyr det at de institusjonelle handlingssvar som hver enkelt kollega utfører, ses i lys av fellesskapets mål. Hvordan forholder forskerne seg til dette?

For det andre er didaktisk rasjonalitet et spørsmål om en plans mulighet for å bli realisert. Vurderinger av dens intensjon og realitet er viktig. Dette er spesielt interessant i forhold til organisasjonsteori: det er ikke bare viktig å se sammenhengen internt i organisasjonens opplæringsprogram. Det er også viktig å kunne vurdere programmets mulighet til å bli aktualisert. En må kunne se sammenhengen mellom den formaliserte forventning som planen viser, og hvordan organisasjonen etterlever det den formulerer om seg selv. Er det laget en godt gjennomarbeidet plan for evalueringen på dette området i evaluering av reform av trosopplæring?

For det tredje handler det om å ha en metarasjonalitet i forhold til en gitt plan. Det innebærer kritiske analyser av læreplanens mål, innhold og metode, av dens konsistens, forankring i virkeligheten og mulighet for å bli realisert, av selve dens grunnlag. Rasjonaliteten viser seg i følge Dale i form av samtaler, i villigheten til å utsette seg for kritikk og i å delta med grunngivelsers for det en sier og gjør. Rasjonaliteten viser seg blant annet i hva som er mulig å kritisere, men også i evnen til forbedringer, i evnen til å korrigere feilslåtte forsøk, når det lykkes i å identifisere mangler og feil (Dale 1998). Dale omtaler her den didaktiske rasjonaliteten i forhold til planer, men jeg mener det også gjelder formaliserte evalueringer.

Evalueringer bør ha de samme krav til rasjonalitet i utførelse som her er nevnt. Hva sier evalueringen av reform av trosopplæring om didaktisk rasjonalitet? Hvilken rolle har forskerne tatt i sitt arbeid? Hva kan leses av refleksivitet ut i fra det de selv uttrykker om eget arbeide?

For å oppnå didaktisk rasjonalitet i forhold til en formulert plan for læring er Dale opptatt av at det må tilrettelegges for kommunikasjon og refleksjon på ulike praksisnivåer i organisasjonen.

### ***Ulike praksisnivåer***

Dale angir skiller i kompetansenivåer i forbindelse med didaktisk rasjonalitet. Han beskriver dem som ulike praksisnivåer i utdanningssystemet, men de kan utmerket godt overføres til ulike praksisnivåer i enhver organisasjon eller virksomhet. Det første kompetansenivået (K1) handler om opplæring i førstelinjetjenesten. Det dreier seg om selve gjennomføringsnivået. Det handler om lærer/elev relasjon i utførsel av en plans formaliserte forventninger. Det andre nivået (K2) handler om utviklingen av undervisningsprogrammer. Her kommer lokal tilpassning/tilrettelegging av planer inn. Overordnede spørsmål på dette nivået er hvilke formålsbestemmelser som skal prioriteres og hvordan den overordnede vedtatte planen kan realiseres. Kommunikasjonen på dette nivået føres mellom kollegaer. Det tredje kompetansenivået/praksisnivået i en organisasjon (K3) dreier seg om kommunikasjon i og konstruksjon av teori. Funksjonen er her å fremføre argumenter, å grunngi standpunkter, og å drøfte begreper. Legitimering blir en viktig praksis. Dette nivået er tilbakeskuende og erfaringsrettet, og må kunne sies å ha et evaluerende fokus.

I evaluering av reform av trosopplæring er evalueringen organisert gjennom de samme nevnte nivåer. De som har sittet nærmest tiltakene har rapportert og gjennom det bidratt til systematiske vurderinger (K1) av hele forsøksprosjektet. De som har hatt ansvar for tilrettelegging av dette arbeidet, har sannsynligvis synliggjort arbeidet i en lagt plan for evalueringen (K2). Evalueringsrapportene som er brukt som grunnlag for denne meta-evalueringen representerer forskernes vurdering av sitt materiale og deres refleksjon over eget arbeid. Her gjøres det også forsøk på å trekke slutninger som kan danne grunnlag for utvikling av teori om trosopplæring (K3).

Konsistens og realisme berører alle tre kompetansenivåene, om enn i forhold til forskjellige spørsmål. Kritiske analyser som drøftinger, angår derimot hovedsaklig det tredje

kompetansenivået, forstått på den måte at en rasjonell revisjon, også der den har relasjon til de to andre kompetansenivåene, økes jo mer dette tredje kompetansenivået i den didaktiske rasjonaliteten (kritisk analyse) blir en del av yrkeskompetansen (Dale 1998). Rom for refleksjon er en forutsetning for den rasjonaliteten Dale legger til grunn. Refleksjonen har igjen best vilkår der det legges til rette for at kommunikasjon og diskusjon rundt faglige spørsmål kan finne sted. Dette forutsetter igjen et felles språk, som er et uttrykk for profesjonalitet og felles faglig ståsted. Hvorvidt det blir lagt til rette for refleksjon i organisasjoner er mye avhengig av ledelse. I krisesituasjoner er handlingsrommet for refleksjon lite og handlingstvangen sterk. Handlingsrommet beskriver Dale som den tiden da en mer eller mindre er i den samme situasjonen; den tiden en har til å overveie forskjellige handlinger. Når handlingsrommet blir utvidet, svekkes handlingstvangen. Når handlingsrommet blir utvidet og handlingstvangen svekkes, øker mulighetene for handlingsrefleksjon (Dale 1998).

Dale viser her at rom for refleksjon er avgjørende for å utøve profesjonalitet i en organisasjon. Er det lagt til rette for refleksjon på ulike kompetansenivåer i evaluering av reform av trosopplæring?

### **3.2.3 Evaluering i "lærende organisasjoner"**

Hvor viktig du ser på deltakelse generelt, hvilke grupper som skal involveres, og hvor mye tid som skal brukes på de forberedende prosessene er avhengig av hvilket syn du har på organisasjonen. I reform av trosopplæring og evalueringen av den er det lagt til grunn "en lærende organisasjon" som organisasjonsteoretiske ståsted. Reformarbeidets forsøksmenigheter blir omtalt som lærende organisasjoner. Evalueringsforskerne bruker derfor betegnelsen i sine rapporter når de omtaler reformarbeidet. Hva det innebærer vil jeg si noe om i det følgende.

Det finnes flere definisjoner på en "lærende organisasjon", men jeg har valgt å benytte Pedler, Boydell and Burgoyne (1989)<sup>1</sup> som har følgende: "An organization which facilitates the learning of all its members and continually transforms itself".

---

<sup>1</sup> Dierkes, Meinolf, Antal, Child & Nonaka (2001): Handbook of Organizational Learning and knowledge. OUP: Philip Sadler: side 416.

Sett i lys av denne definisjonen bør evaluering ha som mål å kunne fungere som et godt redskap for å gjøre læring lettere i virksomheter som har som mål å kontinuerlig omforme seg selv. Hvorvidt det er tilfelle med evaluering av reform av trosopplæring vil jeg drøfte nærmere i kap 5. Det stilles store krav til organisasjonens evne til å lære i ”lærende organisasjoner”. Fordi idealet for organisasjonen er å være i en kontinuerlig lærende prosess, stilles det store krav til organisasjonens tilrettelegging for at så skal skje. Den enkelte medarbeider, som sammen med kollegaer er den viktigste ressursen i organisasjonen, skal alene og i samhandling med medarbeidere sørge for at den kontinuerlige lærende prosessen holdes ved like. Dette holder organisasjonen i stadig utvikling, og er den viktigste suksessfaktoren for virksomheten (Senge 1999). I følge Senge må de organisasjoner som skal lykkes i fremtiden være organisasjoner som oppdager hvordan de skal vekke til live menneskers motivasjon og fremme deres evne til å lære på alle nivå i organisasjonen (Senge 1999:10). Hvilke utfordringer eller barrierer eksisterer som kan hindre dette? Og hva innebærer dette for evalueringens rolle som verktøy i en lærende organisasjon?

Dahler-Larsen presenterer i sin bok ”den rituelle refleksjon” en oversikt over ulike organisasjonsteoriens forhold til evalueringer. Den rasjonelle organisasjonen, den lærende organisasjonen, den politiske organisasjonen og den institusjonelle organisasjonen. Han nevner dem i denne rekkefølgen, og fremhever de siste som de mest interessante. Den rasjonelle organisasjonen forstår evaluering som rasjonell vitenskap der evaluering benyttes som et instrument for å utvikle vitenskap. Den lærende organisasjon ser på evaluering som instrumentell og opplysende, der evalueringens funksjon er å bidra til læring. I den politiske organisasjonen er evalueringens funksjon å være en strategisk og taktisk politisk aktivitet. I den institusjonelle organisasjon ses evalueringen på som en rite i organisasjonen, der funksjonen er å revidere institusjonen. Jeg opplever disse perspektivene som interessante da jeg senere skal vurdere evalueringens rolle i organisasjonen.

Det er med andre ord ulike oppfatninger om hva slags organisasjonsform som er mest relevant i vår tid, og ulike syn på den ideelle organisasjon kan være utslagsgivende for forståelse og tilrettelegging av evalueringers rolle i organisasjonen (Dahler-Larsen 2005). Hvilken organisasjonsforståelse vi har vil så legge grunnen for valg av metode. I en lærende organisasjon er det viktig at medarbeiderne settes i stand til å lære, og da må det legges til rette for at det kan skje. Metoden du velger vil da forsterke eller begrense mulighetene for dette.



### 3.2.4 Valg av metode

Valg av metode henger nøye sammen med intensjon og mål for evalueringen i følge Vedung. I evalueringen av reform av trosopplæring er det benyttet både kvalitative og kvantitative metoder, men det er stort sett kun prosjektledere, barn og foreldre som har vært målgrupper for evalueringsforskerne. Det betyr at det er kun involverte i reformens forsøksfase som har deltatt i evalueringen. Vedung argumenterer sterkt for å innlemme også andre interesserte i en evaluering, ikke minst for bedre å kunne avdekke mulige bieffekter av reformarbeidet. Dette ville for eksempel kunne vært deler av organisasjonen som ikke har deltatt i forsøksarbeidet, frivillige organisasjoner, døpte barn og foreldre som ikke har benyttet seg av tilbudet osv.

Innsteget i utarbeidelsesprosessen er av avgjørende betydning for gjennomføring og utfall, og hvilken metode som velges er derfor vesentlig. Vedung argumenterer for at metodevalg bør ha et sterkere fokus enn det ofte har, og at det burde benyttes ulike metoder til ulikt bruk (Vedung 1998). Han beskriver en rekke ulike modeller i sin taxonomi over evalueringsmodeller, som han deler i tre hovedgrupper: substansmodeller, økonomiske modeller og institusjonelle modeller. Han argumenterer sterkt for å benytte bieffekt modellen, som utgår fra mål men ikke er verken målstyringsmodell eller målløs modell. Bieffekt modellen har fokus på de effekter som ikke var tiltenkt. Han presenterer også en helsystem modell, som passer godt der hele organisasjonen skal evalueres. Sammen med brukermøll og interessant modell hører disse inn under substansmodellene. De økonomiske modellene er orientert om produktivitetsmåling, kostnads effektivitets analyser både i forhold til cost-benefit og nytte perspektiver. De institusjonelle modellene er opptatt av de indre forhold i organisasjonen og referer til kollegabedømming og selvevaluering.

I offentlig sektor har det vært mye fokus på måloppnåelse, og derfor har spørsmålsstillinger og utarbeidelse av evalueringer ofte vært utarbeidet innen den enkelte virksomhet med utgangspunkt i virksomhetens mål. Evalueringen av reform av trosopplæring har oppstilt hypoteser og mål for sin evaluering. Evert Vedung er kritisk til denne praksisen, og om evalueringen med dette utgangspunktet kan gi et sant bilde av virksomheten. Målene vil lede forskerne til funn, og i følge Vedung begrense mulighetene for å avdekke et helt bilde.

Hvilken metode som velges vil være avhengig av hvem forskeren er og hvilken bakgrunn, kompetanse og erfaring hun har. Hvem som skal evaluere er derfor et viktig anliggende for

oppdragsgivere, og vurderinger av dette hører naturlig hjemme i en god forberedelse av en evaluering.

### **3.2.5 Hvem skal evaluere?**

Spørsmålet om hvem som skal få oppdraget med å evaluere bør i følge Vedung vurderes nøye. Han legger vekt på at det bør gjøres grundige overveielser i dette spørsmålet foran hver evaluering. Det er grunner til å la virksomheten selv evaluere, der det er overveiende sannsynlig at nærhet til feltet og god kunnskap om det som skjer er avgjørende for resultatet, og der utenforstående kun er et forstyrrende element i implementering eller læreprosessen i etterkant. Men i de aller fleste tilfeller vil det være positivt at det nettopp tilføres et utenfrablikk i prosessen. Av og til vil det også være høyst påkrevet. I utdanningssektoren har vi erfaringer med evaluerere som ikke har hatt tilstrekkelig uavhengig posisjon i evalueringssituasjonen. Peder Haug gjengir sine erfaringer med evalueringen av R94 og L97 og hevder at den første evalueringen bar preg av å ha evalueringsforskere som satt for nær sine politiske oppdragsgivere. Dette førte til at Forskningsrådet fikk evalueringsoppdraget i evalueringen av L97, nettopp for å sikre en uavhengighet i forhold til de som hadde initiert og bemidlet evalueringen. For sterk eierfølelse til den virksomhet som evalueres, kan hindre objektivitet og kvalitet i prosess og evalueringsresultat. Hvilken rolle har forskerne i reform av trosopplæring hatt, og hva forteller rapporteringene om deres evne til utenfrablikk og objektivitet i vurderingene?

For å sikre at evalueringsforskere blir behandlet med tillit, må det gjøres grundig arbeid på valg av forskere. Dette er i likhet med validitet i funn og resultater viktig for evalueringens kvalitet. Kvalitetssikring og validitet er forbundet med evalueringens rolle, og hvorvidt den vil føre til kunnskapsproduksjon.

## **3.3 Evaluering som kunnskapsproduksjon?**

Evalueringens plass i organisasjonen vil være utslagsgivende for om den får oppmerksomhet eller ikke, hvorvidt den vil føre til læring og derigjennom kunnskapsproduksjon i organisasjonen. Hvordan evalueringen behandles vil igjen være avhengig av den kvalitet den har i utarbeidelse, gjennomføring og resultat. Et dårlig evalueringsprodukt vil sannsynligvis og forhåpentlig ha dårligere tyngde og mulighet for å påvirke i organisasjonen. Kvaliteten vil

være avhengig av gode valg i kriterier og tilrettelegging for læring. Hvorvidt det skal opereres med kriterier i evalueringer er et omstridt tema. Noen mener det ikke er nødvendig. Charles Atkin har for eksempel denne definisjonen av evaluering: "evaluation is the collection and presentation of data summaries for decision-making" (Schriven 1991:156, Vedung 1998: s 28). Vedung sier i tilknytning til sin definisjon at hans posisjon er at evaluering innebærer vurdering, og at evaluering derfor forutsetter minst et verdikriterium, men han "anger inte vilka kriterier denne värderingen skal utgå i från" Vedung (1998: s 29).

Hvis vi forutsetter at målet er å gi læring, og at intensjonen er å lære av de resultater som funnene gir, så er det også vesentlig hva slags læring vi kan trekke ut av den og hvor god kvalitet det er på resultatet. Jeg vil derfor i det neste se nærmere på hva det innebærer ved å gå nærmere inn på kriterievalg og hva teori sier om hva som skal til for at evalueringer skal fungere som virksomhet for læring.

### **3.3.1 Kriterievalg**

Best mulig kvalitet og validitet i data gir tillit til evalueringen, og er nødvendig for at evalueringen skal bli tatt på alvor og ha en funksjon i organisasjonen og derfor fungere som kunnskapsproduksjon. Dette gjelder enten evalueringene støtter seg på kvantitative eller kvalitative data. Det er derfor behov for å utvikle objektive standarder og kriterier i begge tilfeller. Utvikling av kriterier er derfor en av evalueringsforskernes viktigste oppgaver. Et forskningsarbeid gir svar på det du spør etter. I kvalitativ forskning vil hvilke spørsmål som stilles, hvordan de er formulert, og hvem spørsmålene stilles til, være avgjørende for kvaliteten av produktet. Kriteriespørsmålet henger derfor som nevnt nært sammen med valg av metode. En av de viktigste begrunnelsene for valg av metode er å sikre holdbarheten av de påstander som fremmes eller de data som benyttes.

Absolutt sikker viten er ikke lenger målet for vitenskapen. Det er lenge siden vi gav opp idealet om sikker kunnskap og erstattet den med sannsynlighet. I dag er vi mer opptatt av hypotesetesting. Jo flere observasjoner som støtter en hypotese, desto høyere sannsynlighet er det for at den er riktig. Forskerne i reform av trosopplæring har som vi har sett formulert ulike hypoteser som utgangspunkt for sin forskning. Dette er i tråd med et semantisk syn på teori, som I følge Tone Kvernbekk defineres av meningsstrukturer/semantiske strukturer. Teorier omhandler ikke fenomener direkte men selekterte parametre som utgjør abstrakte systemer/modeller av virkeligheten. Hjelpesynteser kan være både teoretisk

baserte/forskningsbaserte, og erfaringsbaserte (Kvernbekk 2005). Evalueringsforskernes bruk av hypoteser og teori/praksis-forhold i evalueringen er interessant i forhold til et pedagogisk ståsted. Det kan virke som at forskerne følger Stortingets ”bottom up”- strategi også i evalueringen. Praksisnivået har utspillet, og hjelpehypotesene dannes for å bekrefte eller avkrefte praksiserfaringene. Ved hjelp av analyser av dette er det så et mål å trekke slutninger som kan danne grunnlag for utvikling av plan for trosopplæring og et teoretisk fundament å utvikle en plan ut i fra.

Jeg oppfatter deler av hypotesedannelsen til evalueringsforskere som et forsøk på å konstruere teori. Teorikonstruksjon i henhold til det semantiske teorisynet har to trinn; fra fenomen til abstrakt system og fra abstrakt system til teoretisk lov/postulat. Hypotesene til forskerne er formulert med utgangspunkt i forventet endring, og er derfor i tråd med målrettet teori.

I en hver teori eller forskning er valg og bortvalg av parametre et kritisk punkt. Operasjonalisering av de begreper som representerer parametrene likeså. Siden evalueringsfeltet er preget av lite forskning på kriteriespørsmålet, er mye av tilretteleggingen og ansvaret for evaluerings validitet og kunnskapsproduksjon lagt på den enkelte evaluerer. Det har derfor vært naturlig å spørre seg om hvordan forskerne i evaluering av reform av trosopplæring har løst valget av kriterier. Evalueringer er selvsagt underlagt de samme krav til validitet som forskning generelt. Det handler om å innhente data som kan måle det evalueringen er ment å måle, det handler om validitet i operasjonalisering av de mål som er lagt, og det handler om å tilrettelegge for at de refleksjoner og den form for læring evalueringen sikter mot skal skje. I følge Vedung innebærer som nevnt evaluering i etterkant å oppdage, kartlegge, og bedømme resultater med det formål å oppnå økt selvrefleksjon, dypere forståelse og gjennom dette ta bedre begrunnede beslutninger. For å oppnå økt selvrefleksjon og dypere forståelse har du gjennomgått prosesser av læring. Jeg vil derfor hevde at et viktig verdikriterium i evalueringer er at de skal føre til læring i en eller annen form. Det er derfor nødvendig å vite hva læring innebærer i forhold til evalueringer.

### **3.3.2 Evaluering som virksomhet for læring**

”At lære betyr at inlade sig på en rejse med usikkert mål. For kende man målet, vilde der ikke være tale om læring” Qvortrup ( 2004). Læring blir av Qvortrup beskrevet som en prosess som i et system, enten det er psykisk eller organisatorisk, er stimulert av ytre

påvirkninger eller indre dynamikker. Ut i fra egne forutsetninger og med henblikk på å opprettholde deg selv, kan du ved hjelp av disse dynamikkene endre egen funksjonsmåte. Dette kan du gjøre på en slik måte at du reagerer annerledes enn før på en ytre påvirkning, samtidig som du sammenligner den med din tidligere reaksjonsmåte.

Han deler læringsprosessen inn i forskjellige former for læring. Disse formene eller nivåene for læring kan hjelpe oss til å se viktigheten av å tilrettelegge i evalueringsprosessen for at læring skal kunne skje. ”Første ordens viten” beskrives som simpel viten. Det handler om faktaviten. Undersøkelser av typen ”hva husker du?”, eller enkle markedsundersøkelser hører inn under disse. Hvis det foretas evalueringer på dette nivået, er det trolig av kontrollhensyn for å sjekke om en gitt kunnskap på et felt er på plass. Evalueringer av denne type er ofte av en multipel-choice variant, mekanistiske og enkle å produsere, og derfor hyppige i bruk som for eksempel enkle møteevalueringer. Dette er summative evalueringer og baserte på kvantitative data. Når det er foretatt evalueringer av denne type i reform av trosopplæring, så er det for eksempel for å få et temperaturmål på hva som har fungert i enkelttiltak lokalt.

”Andre ordens viten” er viten om viten, og derav refleksiv eller situert viten. Dette er refleksjon over den praksis som er utført. På dette nivået innføres læring som metavitene. Som lærende føres man fra det kjente til det ukjente på det forrige nivået, og i andre ordens viten iakttar man den grensen man overskrider i prosessen. For å legge til rette for læring på dette nivået kreves det at det gis anledning til å reflektere over det som foregikk på forrige nivå. Vi har altså beveget oss fra enkle spørreskjemaer til mer reflekterende spørsmålsformuleringer og kvalitative undersøkelser. Eksempler på denne type evaluering i reform av trosopplæring vil være rapporteringene som er gjort av den lokale ledelse for tiltak i forsøksmenighetene. Rapporteringene har gitt ledelsen av tiltakene anledning til å reflektere over den praksis som er funnet sted.

”Tredje ordens viten” beskrives som viten om viten om viten. Viten om forutsetningen for refleksiv viten – forutsetningene for din egne refleksjon over hva du kan. For at en evaluering skal oppnå læring på dette nivået, så er det nødvendig å ha refleksjon over den praksisrefleksjonen som foregikk i forrige vitensposisjon. Ledere av virksomheter vil trolig og forhåpentlig nå dette nivået, men det kan ikke forventes at de fleste som deltar i et evalueringsarbeid gjør det uten at det gjøres grep for at det kan skje. I evaluerings-sammenheng kan dette nivået presenteres via spørsmål som; fungerte evalueringen i følge

hensikten? Gir den svar på det vi var ute etter? Eventuelt hvilke endringer bør gjøres før evalueringsprosessene fortsetter? De ulike rapportene som er skrevet av evalueringsforskerne i reform av trosopplæring underveis i forsøks tiden kan ha bidratt til refleksjon rundt evalueringens kvalitet og utforming. Det forutsettes imidlertid at forskernes rapporter er lest og at funnene er kommunisert og gjort til egne erfaringer av dem som er ment å skulle lære av dem. Det er dette nivået for læring som Qvortrup mener rommer mulighet for omlæring, og derfor bør være et minimums krav om nivåoppnåelse dersom målet med evalueringen er at det skal komme læring ut av den.

”Fjerde ordens viten” betegnes som ”verdens viten” eller ”metasystemisk viten”. Viten om betingelsene for det systemet som skaper betingelser for dine refleksjoner over hva du kan. Dette nivået er nådd når jeg vet de historiske, sosiale og kulturelle betingelsene jeg evaluerer ut i fra. Dette er ikke en individuell, men en kollektiv læringsprosess. Den flytter seg fra det tredimensjonale læringslandskap og ut i en fjerde, en iakttakelsesposisjon der man kan endre landskapets topografi. Denne form for læring foregår når grupper av individer endrer forutsetninger for den læringen som foregår, eller når forutsetningene endres av noen utenfor. Evalueringenes forutsetninger, forankring, karakter og kvalitet i møte med omgivelsene er gjenstand for vurdering på dette nivået, sammen med spørsmål om hva som eventuelt bør endres ved den evalueringen som er gjennomført før neste evaluering. Fungerte denne evalueringen i henhold til hensikten? Også de lærendes opplevelse av hva evalueringen har utrettet for dem og fellesskapet/virksomheten er her gjenstand for refleksjon. Der evalueringene avdekker organisatoriske svakheter og det kreves forandringer av organisatoriske eller kulturelle forutsetninger, så karakteriseres dette som endringer i fjerde ordensposisjon. Kommunikasjonen på dette nivået er rettet mot betingelsene for andre personers læringsatferd, og kan her forstås som den måten evalueringsprosessen legges opp og resultatene eventuelt skal behandles, eller evalueringen blir presentert. Samlingene som er gjort på nasjonalt nivå av lokale prosjektledere og andre faginstanser i reform av trosopplæring kan sies å være arenaer som er ment å bidra til denne formen for læring. Evalueringsforskerne har gjennom deltakelse i disse forumene fått anledning til å presentere sine funn og gjennom tilbakemeldinger av involverte også fått anledning til å se og evaluere seg selv og evalueringsarbeidet.

Et system, enten det er individ, gruppe eller organisasjon, er i følge Qvortrup definert ved forskjellen mellom dem selv og dets omverden. Det finnes ikke systemer uten en sosial

omverden, akkurat som det ikke finnes samfunn eller interaksjonssystemer uten individer. Forandringer i det ene systemet, skaper forandringer i forutsetningene som det andre systemet opererer i. Når et system lærer, så henviser det til "sitt selv". Læring skjer med utgangspunkt i at systemet vil opprettholde seg selv. Når det står overfor en omlæring, må det veien om seg selv, og via en hukommelse som kan vurdere en før og etter tilstand. Det er denne prosessen av sammenligning mellom to tilstander som skaper en forandringsprosess, enten det gjelder individuell eller kollektiv endring (Qvortrup 2004).

Teorien i kapitlet så langt har vist at for at evalueringer skal føre til kunnskapsproduksjon og resultere i læring må det tilrettelegges for at det skal skje. Det er nødvendig å legge til rette for at de som er ment å lære av evalueringen får anledning til å reflektere på ulike nivåer rundt sin egen og virksomhetens læring. For at det skal være mulig må omfattende inkludering i ulike deler av prosessen til. Hvordan tilretteleggingen for refleksjon og læring gjøres vil påvirke gjennomføringen. Jeg vil nå gå inn i hva dette innebærer, og hvordan man ivaretar de ulike fasene av en evaluering best mulig i organisasjonen.

### 3.4 Hvordan gjennomføre en evaluering?

Vedung presenterer i sin bok "Utvärdering i politik och förvaltning" to ulike modeller som kan fungere som redskap for å se evalueringens rolle i en virksomhet. Han mener at for å pedagogisk illustrere ideen om et "framåtsyftande tillbakablickanda" bruker den offentlige politikken å betraktes som en syklisk prosess. Vedung mener at en mer fruktbar illustrasjon for intervensjonsprosesser og evaluering i offentlig sektor er å se intervensjon og evaluering som to forskjellige prosesser og derfor som to sjikt i en modell. Jeg vil videre beskrive de ulike kategoriene i flerfasemodellen, og sette det i sammenheng med det vi har sett av teori om hvordan en evaluering bør utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i en organisasjon.

#### *Flerfasemodellen*

Den offentlige intervensjonsprosessen innledes med initiering. Noen identifiserer et problem og iverksetter tiltak. Deretter foregår det Vedung kaller "beredning". Problemet bearbejdes og en prosess forberedes ved for eksempel innsamling av informasjon, utvikling av

målformuleringer, og kartlegging av økonomiske konsekvenser. Dette er kanskje den viktigste fasen i en intervensjonsprosess, der forutsetningene for resultatet legges.

Etter at nødvendig forberedelse eller ”beredning” er gjort tas en beslutning. En formell legitimerende avgjørelse faller om valg av strategier for handling. Dette følges av selve gjennomføringsprosessen eller implementeringen som omhandler både forvaltning og sluttprestasjoner. Forvaltningen i reform av trosopplæring er ledelsen i tiltakene og de som faktisk utfører intervensjonens innhold. De resultatene som kommer ut av tiltakene kaller Vedung sluttprestasjoner. Disse prestasjonene oppsummeres i det han kaller ”utfall”, som betegner de effekter som intervensjonen har gitt. Når dette punktet er nådd i en intervensjonsprosess bør det iverksettes en etterkontroll i form av en evaluering.

Evaluering er i følge Vedung en systematisk undersøkelse og bedømmelse av sluttfasene i en intervensjonsprosess i offentlig sektor, der evalueringens rolle er av det retrospektive og tilbakeskuende slaget. Den tilbakeskuende rollen forbindes med framtiden gjennom en ny start på intervensjonsforløpet med ”återkopplingsmekanismen”. Ved å ta et skritt tilbake og ut av evalueringsprosessen og vurdere alle trinn fra utsiden, vil det være mulig å skape et bedre grunnlag for å gå videre og for å ta ”betre grundade beslut”. Återkopplingsprosessen er prospektiv og framtidsrettet.

Poenget med flerfasemodellen til Vedung er å skape en mulighet for å vurdere den gjennomførte prosessen fra initiering til utfall før endringer og utbedringer iverksettes. Flerfasemodellen ivaretar både det tilbakeskuende perspektivet og refleksjon over det som har skjedd eller skjer, samtidig med framtidige virkninger på de neste beslutninger.

Flerfasemodellen til Vedung opptar meg fordi den gir et godt bilde av de ulike fasene som en intervensjon eller en reform i offentlig sektor inneholder. Den viser også viktigheten av at evaluering både er en viktig del av en intervensjon samtidig som den må foregå litt utenfor intervensjonen for å ha den riktige og nødvendige rolle i den helhetlige prosessen.

Det er grunn til å dvele ved den delen av intervensjonsprosessen som Vedung har kalt ”beredning”. Jeg ser på denne fasen som en viktig didaktisk fase, der grunnlaget for kvaliteten på det forventede resultatet legges. Også forberedelsen av hvordan intervensjonen skal evalueres bør etter mitt syn gjøres i denne fasen. For å få gode og målbare resultater av evalueringer er det nødvendig å forberede evalueringsarbeidet planmessig før man går videre i prosessen.



---

For å belyse denne sammenhengen har jeg utarbeidet en egen modell som jeg vil presentere i neste kapittel. Modellen har jeg kalt: didaktisk evalueringsmodell.

### *Oppsummering*

Dette kapitlet har handlet om å belyse problemstillingen i oppgaven i forhold til teori. Jeg har presentert ulike innfallsvinkler til dette ved hjelp av Evert Vedungs teori om hvordan en evaluering helst bør tilrettelegges, hvordan konsistens i en formulert plan for læring og didaktisk rasjonalitet influerer på ulike kompetansenivåer gjennom Erling Lars Dales teori, og hvordan man best kan sikre refleksjon og læring gjennom evalueringer ved hjelp av Lars Qvortrups vitensposisjoner.

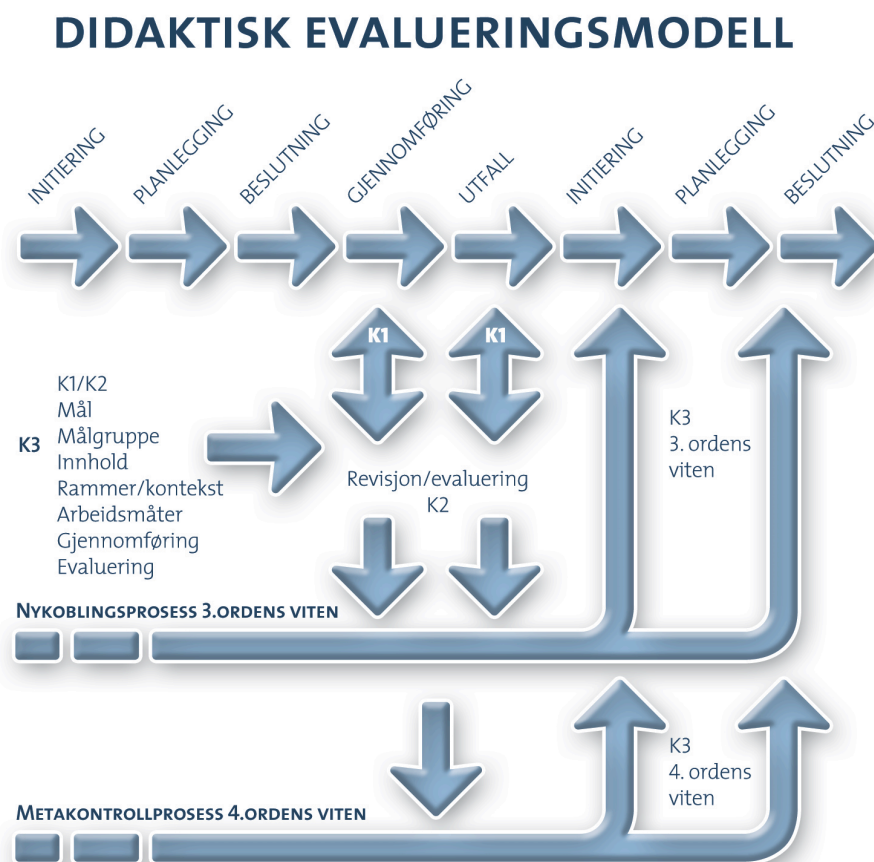
Gjennom arbeidet med Evert Vedungs flerfasemodell, og de nevnte teoretikeres teorier om didaktikk og læring, har jeg funnet fram til en modell som jeg har kalt en didaktisk evalueringsmodell. Modellen forener, slik jeg ser det, Vedungs engasjement for de ulike fasene i en evalueringsprosess i offentlig sektor, og mitt viktigste anliggende som jeg beskriver gjennom den didaktiske relasjonsmodellen og Dale og Qvortrups teorier for tilrettelegging av refleksjon og læring; viktigheten av didaktisk arbeid i forberedelse av evalueringsarbeider. En prospektiv holdning og innfallsvinkel til evalueringsfeltet, vil kunne bidra til bedre kvalitet i evalueringsresultater og føre til bedre begrunnede beslutninger i offentlig sektor.

Hva modellen gir av kunnskap om evaluering i reform av trosopplæring vil jeg beskrive nærmere i kapittel 5. Men først vil jeg presentere den didaktiske evalueringsmodellen i lys av teori om didaktikk og evaluering.

## 4. Den didaktiske evalueringsmodellen

I dette kapitlet vil jeg presentere modellen jeg har utviklet gjennom arbeidet med oppgaven. Modellen er som nevnt basert på flerfasemodellen til Evert Vedung, og de didaktiske kategoriene i den didaktiske relasjonsmodellen til Bjørndal og Lieberg (1978). Dales teori om didaktisk rasjonalitet og hans ulike kompetansenivåer, samt Qvortrups vitensposisjoner er benyttet i modellen for å synliggjøre hva tilrettelegging for læring innebærer i de ulike fasene i en evaluering. Jeg vil i det følgende beskrive hva jeg mener modellen viser, og hva jeg mener å tilføre av teori gjennom utviklingen av modellen. Ved å beskrive modellen mener jeg samtidig å gi svar på problemstillingen: **Hvordan bør en evaluering utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”?**

Modellen jeg har utarbeidet ser slik ut:



Jeg vil benytte den didaktiske evalueringsmodellen som utgangspunkt for drøftingene i

analysen i kapittel 5. Jeg vil derfor komme tilbake til hvordan jeg mener evalueringen av reform av trosopplæring har forholdt seg til de ulike fasene i den didaktiske evalueringsmodellen, og hvordan evalueringen eventuelt burde vært tilrettelagt for å gi et bedre læringsutbytte enn den har gitt. I dette kapitlet vil jeg kun beskrive modellen i lys av teorien fra kapittel 3.

## 4.1 Hvorfor en didaktisk evalueringsmodell?

Flerfasemodellen til Vedung viser hvilke faser en evaluering i offentlig sektor bør forholde seg til for å ivareta helheten i et evalueringsarbeid, og sørge for at den evaluering som blir utarbeidet og gjennomført virker etter sin hensikt (Vedung 1998). Vedungs modell viser derfor gangen i en evalueringsprosess, men gir ingen nærmere bestemmelse av hva de ulike fasene innebærer. Vedungs beskrivelser av fasene er nokså generelle, og jeg oppfatter at han nettopp mener det er et poeng ikke å angi noen verdikriterier spesielt. Han mener det bør være minimum ett verdikriterium i evaluering, men han vil ikke bestemme hvilket. Mitt arbeid i denne oppgaven har nettopp vært å forsøke å finne fram til verdikriterier. Jeg mener at jeg ved hjelp av en didaktisk tilnærming til evalueringsprosessen, både gir en nærmere bestemmelse av hva fasene innbærer, og derigjennom samtidig bidrar til å definere verdikriterier for evalueringen. Dette vil jeg forklare nærmere gjennom kapitlet.

Den didaktiske relasjonsmodellen som ble presentert i kap 3.2.2 viser at evaluering er en naturlig del av didaktisk virksomhet. Didaktiske overveielser og avgjørelser er preget av både vår kunnskap om undervisnings- og oppdragelsessituasjoner, og de intensjoner vi har for vår didaktiske virksomhet (Engelsen 1990).

Et vesentlig trekk ved didaktikken er dens forsøk på å begrepsfeste sitt studiefelt ved hjelp av de didaktiske kategorier. Jeg har lagt de didaktiske kategoriene inn i den didaktiske evalueringsmodellen under Vedungs fase "beredning" som jeg har omdøpt til "planlegging". Scwab (1983) kaller slike kategorier for "the commonplaces of curriculum" (Engelsen 1993). Kategoriene er forhold som man i følge Engelsen er nødt til å ta stilling til ved enhver

gjennomtenkning og planlegging av undervisning eller program for læring. De er derfor også gjeldende for planlegging av evalueringer og i arbeidet med å utvikle evalueringsprogram<sup>2</sup>.

Ved hjelp av et program vil evalueringen kunne utarbeides slik at den kan fungere som et best mulig grunnlag for læring. Vedungs åttepunkt-program for ledelse av evaluering gir viktige innspill til hva et evalueringsprogram bør innholde, og programmets ulike betegnelser fungerer i modellen som navn på de ulike fasene evalueringen i sin helhet inneholder. Didaktisk rasjonalitet i et utviklet program for opplæring, er i følge Dale et spørsmål om konsistens. Det innebærer at det skal være en sammenhengende relasjon mellom planens mål, innhold og metodevalg.

Enten evalueringen handler om en mindre del av organisasjonens virksomhet, eller den er ment å evaluere organisasjonen som helhet, handler tilretteleggingen for læring i en ”lærende organisasjon” om å gi alle medlemmer av organisasjonen mulighet for refleksjon og deltakelse i evalueringen. Modellen viser hva tilretteleggingen innebærer i henhold til Dales ulike kompetansenivåer og Qvortrups vitensposisjoner. Selve initieringen av en evaluering i en organisasjon krever refleksjon på K3 nivået i følge Dale, fordi det forutsetter refleksjon over evalueringens plass i organisasjonen og ikke bare fokus på selve utviklingen av et program for opplæring. Vi befinner oss derfor også på tredje og fjerde ordens viten i henhold til Qvortrups vitensposisjoner. For å sikre en god kultur for læring og at alle som er ment å lære av evalueringen er gitt mulighet for det, forutsettes det involvering også i starten av en evaluering, og i utviklingen av et program for den evaluering som skal skje. Jeg har av den grunn satt K3 inn under initiering i modellen.

De ulike fasene i en evalueringsprosess har hver for seg egne problemstillinger knyttet til kvalitetssikring av evalueringsarbeidet. En del av denne kvalitetssikringen handler også om å angi retning for hele evalueringsprosessen slik at alle ledd bidrar til læring. Jeg vil i det følgende gå gjennom hver fase i den didaktiske evalueringsmodellen og begrunne hvorfor jeg mener den bidrar med nye perspektiver i evalueringsforskning.

---

<sup>2</sup> Dale bruker begrepet program som et videre og mer samlende begrep enn plan, der programmet sees som en del av organisasjonens formaliserte forventninger. I en læringsorganisasjon er planlegging bare en delaktivitet i forhold til organisasjonens mer samlende program (Dale 1998). Jeg vil videre benytte program som betegnelse for planen som helhet, og mener begrepet program fungerer bedre også i forhold til det didaktiske og formaliserte arbeidet med evalueringer i organisasjoner.

#### 4.1.1 Forberedelsesfasen

Den didaktiske evalueringsmodellen viser innholdet i en didaktisk tilnærming til evalueringer først og fremst gjennom forberedelsesfasen. Modellen viser forberedelsesfasen gjennom kategoriene initiering og planlegging (Vedungs ”beredning”). Forberedelse innebærer planlegging, og planlegge er å forhåndsbestemme (Dale 1998). Det handler om å bestemme *hva* som skal gjøres, *hvordan* en skal oppnå det ønskede resultat, hvilket *tidsrom* planen er ment å gjelde for, og *hvem* som skal gjøre hva til hvilket tidsrom. Planleggingens funksjonsverdi er å være organisasjonens grunnlag for å utvikle undervisningsprogram (Dale 1998). Med undervisningsprogram menes all undervisning eller opplæring som er ment å føre til læring og forandring hos medlemmene i organisasjonen, i så måte også evalueringer.

Starten på en evaluering danner på flere måter grunnlaget for både gjennomføring og resultat. Det er i denne fasen alle føringer blir lagt, og alle rammer og forutsetninger skapes. Disse får igjen konsekvenser for alle de neste leddene i en evalueringsprosess. Det er med andre ord i denne fasen at selve grunnlaget for læring legges. Dale vektlegger behovet for refleksjon og handlingsrom i utførelse for å sikre best mulig læringsutbytte. Målet er at alle medlemmer og aktører i en ”lærende organisasjon” skal oppnå alle tre kompetansenivåer, og dette krever at den planen som utvikles gir rom for at dette kan skje. Qvortrup viser gjennom sine vitensposisjoner at ulik form for læring krever ulik tilrettelegging for at læring skal finne sted. I utarbeidelsen av et program for evaluering vil forskerne og de involverte ha behov for å tilrettelegge for læring på første og andre ordens viten, uten å gi slipp på de overordnede intensjoner og mål for evalueringen som helhet i og utenfor organisasjonen, som igjen krever tilrettelegging for refleksjon på tredje og fjerde ordens viten. Det handler om å finne gode grep for å presentere faktisk viten fra førstelinjetjenesten og sette dette inn i en situert kontekst, og i en institusjonell og samfunnsmessig helhet, på en måte som gjør at alle medlemmer av organisasjonen kan lære av det. Jeg vil vise i gjennomgangen av forberedelsesfasen hva denne tilretteleggingen innebærer. Forberedelsesfasen deles inn i initiering og planlegging. Jeg vil i det følgende beskrive forhold som har betydning for utarbeidelse av planen i denne fasen. Initiering vil i hovedsak dreie seg om forhold knyttet til formål og intensjon ved evalueringen, og hvordan planen må tilrettelegges for å ivareta dette. Under planlegging skal alle forhåndsbestemmelser operasjonaliseres.

## *Initiering*

Modellen viser at evalueringer starter med en idé om å evaluere som Vedung kaller initiering. Selve ideen sier noe om forventninger til evalueringen, som det vil være viktig å kjenne til når de nødvendige forberedelser skal gjøres. Det handler om å tydeliggjøre evalueringsprogrammet i forhold til formål og intensjon. Formål er overordnede verdipremisser, angitt i en overordnet plan (Dale 1998). Formålet befinner seg derfor på K3-nivået i henhold til Dales kompetansenivåer, og innebærer som vist å framføre argumenter, grunnleggende standpunkter, og drøfte begreper. Legitimering blir en viktig praksis. Dette planleggingsarbeidet forutsetter det nivået for læring som Qvortrup beskriver som fjerde ordens viten. Forhold knyttet til programmets rolle og innflytelse i og utenfor organisasjonen, samt visjoner for organisering og læring i organisasjonen og hva det innebærer for en "lærende organisasjon", hører inn under refleksjonen som bør foregå i denne fasen.

## **Evalueringens formål**

Som start på en evalueringsprosess hører det også hjemme å bestemme mål for evalueringen. Overordnede mål for evalueringen i tråd med evalueringens formål og hensikt, hører med under den didaktiske evalueringsmodellens første fase initiering. Formålet med evalueringen skal operasjonaliseres i mål som skal danne grunnlag for en felles grunn å arbeide ut fra for alle medlemmene i organisasjonen.

Dale har vist at å gjennomføre opplæring med didaktisk rasjonalitet innebærer målrettet læring, og gjennom det, felles referanseramme og forståelseshorisont. I en organisasjon samordner kollegene forskjellige kulturelt bestemte handlingssvar. Samarbeid, teamwork, betyr at enhver kollega har identifisert seg med fellesskapet, samtidig som deltakerne utfører ulike funksjoner innenfor det. Gruppen blir en organisert enhet om medlemmene utfører de ulike funksjonene i lys av felles mål (Dale 1998). Med referanse til Mead bruker Dale begrepet "den generaliserte andre" når han beskriver holdningen til det kollegiale fellesskapet som helhet. Kollegene ser seg selv som medlemmer i organisasjonen i det de inntar den generaliserte holdningen til seg selv. De handler da på vegne av et fellesskap. Den profesjonelle aktøren handler i lys av ønsker som er forandret til bevisste mål. Aktørens frihet i gjennomføringen av en opplæring ligger ikke i å handle etter tilfeldige tilbøyeligheter. De umiddelbare handlingene i en opplæring er mer å forstå som en

---

manifestasjon av en omfattende og sammenhengende plan som aktøren ønsker å realisere (Dale 1998).

Behovet for å samles om felles mål er derfor viktig innenfor all virksomhet som har læring som mål, og spesielt i ”lærende organisasjoner” som har som mål å legge til rette for læring hos det enkelte medlem/aktør. Det er derfor også et viktig anliggende for evalueringer. For å ivareta didaktisk rasjonalitet i utførelse, har Dale påpekt betydningen av konsistens. Dette innebærer også at eksterne forskere som kommer utenfra må forholde seg til de samme mål som organisasjonen styrer etter, hvis evalueringen skal fungere rasjonelt og fungere overfor medlemmene av organisasjonen. Evalueringer kan få betydning som en medierende artefakt (Säljö 2006) overfor den enkelte profesjonelle aktør dersom den blir en del av deres felles praksis, og evaluerer denne i lys av felles mål. På denne måten vil evalueringens læringseffekt ivaretas best mulig. Den vil også på denne måten få betydning for bygging av kultur i organisasjonen som helhet, fordi den vil bli et felles eie og et felles verktøy for læring og utvikling.

Initieringen har derfor også følger for evalueringens rolle i organisasjonen. Alle forutsetninger og rammer som er nevnt så langt i dette kapitlet legger grunnlaget for den læring som evalueringen er ment å gi. Svikt i denne fasen av et evalueringsarbeid skaper et lite heldig utgangspunkt for det videre arbeidet i evalueringen. Det er denne delen av forberedelsesfasen som skal danne grunnlaget for den videre utarbeidelsen og tilretteleggingen av evalueringen. Neste fase i forberedelsesarbeidet er planlegging.

### *Planlegging*

Når ideen om å evaluere er skapt, og intensjon ved den er identifisert og de ytre rammene rundt den er på plass, starter arbeidet med å tilrettelegge. I denne sammenhengen handler det om i forberedelsesfasen å utarbeide programmet for evalueringen på en slik måte at evalueringen kan gi et best mulig grunnlag for læring i en ”lærende organisasjon”. Det skal utvikles verktøy, og organiseres ressurser, slik at intensjon og mål kan nås. Det er altså med utgangspunkt i forrige fases arbeid at den videre forberedelse gjøres. Operasjonalisering av mål og valg av metoder for å nå målene hører derfor inn under denne fasen. Vi befinner oss derfor, slik den didaktiske evalueringsmodellen viser, på K1 og K2-nivået jamfør Dales kompetansenivåer. I denne delen av programutviklingen kreves det refleksjon over hvordan planleggingen i initieringsfasen kan operasjonaliseres og tilrettelegges, slik at det sørges for

en best mulig førstelinjetjeneste der mål og intensjoner ivaretas. Dette gjøres best gjennom å utvikle et program for evaluering som legger opp til gode refleksjonsmuligheter over den praksis som foregår i førstelinje nivået K1.

Punktene som er nevnt innledningsvis kan også beskrives som de didaktiske kategorier. Mål, Målgruppe, Innhold, Rammer/kontekst, Arbeidsmåter, Gjennomføring, og Evaluering hører inn under denne betegnelsen. De didaktiske kategoriene bidrar til å definere kriterier i evalueringen ved å angi de standarder evalueringen skal mål etter, og jeg mener derfor de også bidrar til å bestemme verdikriterier. Jeg vil i det følgende gå inn i de ulike kategoriene og kort beskrive hva teori sier om innholdet i dem. De valg som tas i kategoriene vil stå i et særlig forhold til hverandre, slik Bjørndal og Lieberg (1978) har beskrevet det i forhold til den didaktiske relasjonsmodellen.

### **Operasjonalisering av Mål**

De overordnede målsettinger som ble dannet under initiering i den didaktiske evalueringsmodellen skal i denne fasen operasjonaliseres og beskrives i et program for evalueringen. Jeg har under initiering omtalt viktigheten av felles mål i relasjon til Dales begrep ”didaktisk rasjonalitet”, og betydningen av et overordnet mål som bekrefter formålet med en virksomhet. I forbindelse med dette har jeg også forsøkt å vise til betydningen av at både eksterne og interne aktører i et evalueringsarbeid må forholde seg til den samme vektning av rasjonalitet og konsistens, og derfor velge å forholde seg til samme mål. Jeg ser dette som forutsetning for at en evaluering i utarbeidelse og gjennomføring skal kunne gi et best mulig grunnlag for læring i en ”lærende organisasjon”.

Det skal utledes mål av den overordnede målformuleringen i denne fasen. Denne prosessen handler om å formulere mål som kan operasjonaliseres og derigjennom besvares gjennom evalueringen. Disse operasjonaliseringene må gjøres med det utgangspunkt at de svar man er ute etter vil føre til økt selvrefleksjon, dypere forståelse, slik at man gjennom dette kan ta bedre begrunnede beslutninger (Vedung 1998). Operasjonalisering av de mål som er lagt er av det viktigste arbeidet i planleggingsfasen. I en hver forskning er valg og bortvalg av parametre et kritisk punkt. Operasjonalisering av de begreper som representerer parametrene likeså.

Tilliten til evalueringer som grunnlag for læring i en organisasjon og overfor den enkelte aktør i organisasjonen er avhengig av kvaliteten på det materialet som hentes inn, og om det



oppleves av organisasjonens medlemmer å være konsistens og sammenheng mellom mål og de operasjonaliseringer som er gjort. Dette gjelder også det materialet som er utarbeidet for å gi svar på om målene er nådd. Aktørenes selvrespekt knyttes til å realisere felles mål (Dale 1998). I "en lærende organisasjon", der medlemmene er ment å være sentrum for den lærende virksomheten, er det derfor avgjørende at arbeidet med målformuleringer og operasjonalisering av de mål som er lagt, har høy kvalitet. Målene skal videre formuleres på en måte som sikrer at læring skjer, og at aktørene forandres. De skal si noe om hvilken betydning innholdet er ment å ha for aktørene, og formål og grunnleggende verdier må kunne spesifiseres og gis operasjonell funksjon (Dale 1998: Tyler 1949)

### **Målgruppe**

Å definere målgruppe er en nødvendig forutsetning for å tilrettelegge for læring. Hvem evalueringen henvender seg til har betydning for hvordan den bør utarbeides og gjennomføres. Det handler om å definere hvem denne målgruppen er, og hva den krever av tilrettelegging når det gjelder språk og kommunikasjon, utvikling av rapporteringsverktøy, presentasjon av evalueringen, hvem som skal involveres i alle deler av evalueringsprosessen, samt hvordan dette best bør gjøres. Hvordan man lykkes med den tilretteleggingen som valg av målgruppe krever, vil igjen ha betydning for hvorvidt evalueringen vil føre til læring. Hvis målet med evalueringen er at den skal fungere som et refleksjonsgrunnlag for medarbeidere i "en lærende organisasjon", betyr det at målgruppen i realiteten er hver medarbeider i organisasjonen. Dette stiller store krav til tilrettelegging og planlegging, samt hvordan programmet for evalueringen må utarbeides. Dette må møtes på en helt annen måte enn hvis målgruppen er begrenset til ledelse og oppdragsgivere.

I evalueringssammenheng har målgruppedefineringen to formål. Den skal både definere hvilken målgruppe som er ment å skulle lære av evalueringen, og hvem evalueringens objekt er. Dette vil ofte være samme målgruppe, men trenger ikke å være det. I tillegg kan evalueringen ha som mål å nå målgrupper som ikke er direkte involverte i evalueringens innhold, men likevel være ønsket av for eksempel oppdragsgivere som målgruppe for å lære av evalueringen.

### **Innhold**

Det neste som skal defineres er innholdet i evalueringen. Hva skal evalueringen evaluere? Hva skal evalueringen orienteres om? Identifisering av dette i forhold til intensjon og mål

med evalueringen, handler om å innholdsbestemme. Det overordnede spørsmålet er her slags tilrettelegging og forberedelse som kreves i forbindelse med innholdsbestemmelsen for at evalueringen skal føre til læring.

Innholdet i evalueringer vil variere fra evaluering til evaluering, og kan være alt fra omfattende og omfangsrike evalueringer innholdsmessig, som iverksettes for å vurdere alle sider ved en organisasjons virksomhet, til evalueringer av mindre prosjekter. Like fullt kreves det en nøye spesifisering og definering av hva som skal evalueres, vurderes og måles, samt hvilken målestokk/standard som bør velges som utgangspunkt for denne målingen. Dette skal brukes som utgangspunkt for valg av arbeidsmåter og metoder. Innholdet henger nøye sammen med målet, og som nevnt under mål er det i følge Dale en forutsetning for rasjonalitet i utførelse at medlemmer og ledelse av en organisasjon har den samme forståelseshorisonten. Avklaring av innholdet som evalueringen skal orienteres om er i følge den didaktiske evalueringsmodellen nødvendig for å skape felles forståelse hos medlemmene i organisasjonen. Refleksjonen (K2) hos den enkelte medarbeider, og mellom kollegaer i organisasjonen i selve gjennomføringsfasen, bør ikke handle om hva slags innhold som skal evalueres, eller med hvilken standard innholdet skal måles/evalueres, men heller hvordan evalueringen skal utføres, presenteres, og behandles i organisasjonen. Innholdsbestemmelse og valg av kriterier for hva evalueringen skal måle/evaluere må gjøres i planleggingsfasen.

### **Rammer/kontekst**

Å bestemme rammefaktorer innebærer å definere hva som trengs av ressurser for å iverksette og gjennomføre evalueringen i tråd med mål, målgruppe, innhold og de krav som standarder og kriterievalg tilsier. Uten at rammene gir rom for utarbeidelse og gjennomførelse i tråd med dette, kan ikke evalueringen utføres med kvalitet og didaktisk rasjonalitet (Dale 2006). Det handler om alt fra teknisk verktøy til innhenting av data som skal benyttes, til hvilke personalressurser og eventuelt reisekostnader som trengs for å gjennomføre hele evalueringen.

De økonomiske rammene som oppdragsgiverne tildeler evalueringsformålet vil begrense eller gi rom for kvalitet i alle deler av prosessen. Forskernes oppgave blir her å definere behov i forhold til oppdraget, og gi tilkjenne de rom eller begrensninger som rammene gir. Siden kostnader knyttet til evalueringer er avhengig av oppdragsgivernes intensjon og behov vedrørende bruk, vil det være nødvendig med tilpassninger i forhold til hver enkelt

evaluering. Det må innebære at forskere også må anse det som sin oppgave å påvirke og informere om hvilke muligheter eller konsekvenser en kostnadsramme vil ha.

Valg av kontekst hos forskerne, har igjen følger for hvordan tilretteleggingen av læring skal skje. Ulike kontekster krever ulik tilrettelegging for læring, og tilsier endringer i språk, formidling, og organisering av evalueringen. I følge Qvortrup innebærer tilretteleggingen at medlemmer av organisasjonen må ha rom for ulik type refleksjon ut fra hvilke spørsmålsstillinger og typer evalueringer det handler om. Summative og formative evalueringer krever forskjellig type refleksjon for at de skal føre til læring. Enkle markedsundersøkelser eller innhenting av data for oppslutning av et tiltak er eksempler på summative evalueringer som er baserte på kvantitative data. Disse forutsetter simpel eller faktisk viten hos de som er ment å lære av dem (Qvortrup 2004). Rammefaktorene vil bestemme kvaliteten på det materialet som disse undersøkelsene bygger på. Kvaliteten vil igjen ha betydning for bruken av evalueringen som utgangspunkt for refleksjon og læring hos den enkelte medarbeider. Om denne refleksjonen finner sted vil igjen ha betydning for om evalueringen vil bidra til læring i ”den lærende organisasjonen”.

Formative evalueringer er basert på refleksiv eller situert viten, som krever refleksjon over den praksis som er utført. Disse formene for evalueringer vil kreve større grad av rom for refleksjon. I ”en lærende organisasjon” krever tilrettelegging av dette omfattende arbeid slik at alle medarbeidere i organisasjonen blir inkludert og deltar. Valg av rammefaktorer vil ha følger for hvilke rom for refleksjon hver enkelt medarbeider gis, slik at de kan utvikle profesjonalitet i arbeidet (Dale 1998). Rammefaktorene som velges vil begrense eller innby til inkludering og deltakelse, og derigjennom læring i hele organisasjonen. Uansett må de valg som her gjøres, følges opp i gode valg av arbeidsmåter om de skal gi resultater.

### **Arbeidsmåter**

Valg av arbeidsmåter handler om å velge rett metode til de behovene som er kommet fram i arbeidet med de foregående kategoriene. Hvilke arbeidsmåter skal velges for å oppnå målene, intensjonen med evalueringen? Hvilke metoder fungerer best overfor målgruppen, og hvilke metoder egner seg best for å måle/evaluere det innholdet som er definert ut i fra de standarder som er gitt? Vedung argumenterer for mer fokus på metodevalg og at evalueringsforskere i sterkere grad bør benytte ulike metoder til ulikt bruk. Dette forutsetter omfattende planleggingsarbeid i forhold til det enkelte evalueringsoppdrag. Det er likevel

grunn til å dvele ved dette, ikke minst med tanke på å oppfylle krav til validitet og utførelse av evalueringer i tråd med mål og de foregående nevnte kategorier, samt de kriterievalg som er kommet ut av arbeidet med dem.

Vedungs taxonomi over ulike evalueringsmodeller inneholder mange ulike metoder som hver for seg og samlet dekker et bredt spekter av måter å evaluere på. De gir alle ulike innfallsvinkler til evalueringsoppdrag. Dette innebærer igjen at hvis flere av metodene tas i bruk i en evaluering, sikres flere perspektiver plass i den samlede evalueringen, som igjen kan bidra til bedre validitet i resultat og gjennomføring. Jeg vil i anvendelsen av den didaktiske evalueringsmodellen i neste kapittel gå nærmer igjennom de metodene som Vedung beskriver i sin taxonomi, for å vise hva de kunne tilført evaluering av reform av trosopplæring av perspektiver og kvalitet i utførelse hvis de var benyttet. Her vil jeg nøye meg med å si meg enig med Vedung i at bruk av flere metoder, eller valg av bieffekt-metode som gir en større bredde i fokus enn for eksempel måloppfyllelsesmodeller, vil ha stor betydning for kvalitet av evalueringer.

Metodevalget må i alle tilfelle ha som mål å underbygge og bidra til å gi svar på de kriterier som er bestemt i arbeidet med de foregående kategoriene i dette kapitlet. Som del av dette arbeidet skal også valg av arbeidsmåter og metoder bidra til å tilrettelegge for hvordan det som er utarbeidet i de andre kategoriene skal læres.

## **Gjennomføring**

Hvis planleggingen skal bli god nok må det også tenkes gjennom hvordan evalueringen vil fungere i en gjennomføringsfase. Det holder altså ikke å legge planer for så å sette de direkte ut i livet. Det må gjøres en forutgående vurdering av hvordan planen vil fungere i praksis. Denne kategorien ivaretar dette fokuset, som handler om en tenkt gjennomføring. Dette er på mange måter planlegging av hvordan evalueringen helst eller ideelt sett bør gjennomføres. Ved å ta dette på alvor er muligheten tilstede for justeringer før iverksetting. Det handler om å se for seg hvilke utfordringer som kan dukke opp, og eventuelt hvilke forhåndsregler som må tas for å håndtere dem. Det handler også om å ha fokus på hva man forventer å oppnå, og hva som trengs av ressurser og tilrettelegging. Er det for eksempel behov for støttefunksjoner?

I løpet av arbeidet med denne kategorien blir det lagt føringer for hvordan gjennomføringsfasen vil komme til å bli. Det som ikke er planlagt vil sannsynligvis ikke inntreffe, og det

som er planlagt vil kunne gi uventede effekter. Det er et mål å unngå det sistnevnte da det vil kunne avdekkes med bedre fokus på planleggingen av denne fasen. Jeg vil senere i oppgaven komme tilbake til selve gjennomføringsfasen i evalueringsprosessen. I denne delen av oppgaven er derfor fokuset på hva rapporteringene sier om planleggingen av gjennomføringsfasen. Det er likevel nok en grunn for at evaluering ligger som en didaktisk kategori i forberedelsesfasen under planlegging. Også den evaluering som er gjennomført bør revideres og ved det sees i sammenheng med de andre didaktiske kategoriene når evalueringen revideres. Denne prosessen skjer i gjennomføringsfasen og vil derfor bli omtalt mer under punkt 4.1.2 i oppgaven.

Det kan selvsagt oppstå ting underveis i en evaluering som ikke er mulig å forutse uansett hvor godt du forbereder og planlegger. Likevel vil jeg argumentere for at det er lite læring som kommer ut av et program som ikke er tilrettelagt for at læring kan skje. Til dette hører planlegging. Det handler derfor i arbeidet med denne kategorien om å vurdere og reflektere over (K3) hvordan læringen som er planlagt å skje (K2) skal utføres (K1), slik at den vil gi et læringsutbytte. Dette handler igjen om i planleggingen å legge til rette for refleksjon hos det enkelte medlem i den ”lærende organisasjonen” i tråd med Qvortrups vitensposisjoner.

### **Planlegging av evaluering**

Evaluerer er en naturlig del av tilrettelegging av læringstiltak i henhold til den didaktiske relasjonsmodellen som ble presentert i kap. 3.2.2. I den didaktiske evalueringsmodellen er evaluering plassert innenfor de didaktiske kategoriene under planlegging i forberedelsesfasen. Dette mener jeg sier noe om hvilken rolle evalueringer skal ha for å evne å føre til læring. Den må både ha en rolle innenfor reformens virksomhet og samtidig stå trygt utenfor og bidra med utenfrablikk. Dette forutsetter god planlegging av evalueringen før den iverksettes, slik at begge roller ivaretas og tilrettelegges for. Dette gjelder også følgeforskning fordi evalueringens rolle er den samme på dette punkt enten den er summativ eller formativ. Det handler om hvordan evalueringer skal organiseres for å ivareta kvalitet og læringsutbytte, og da kreves det at den både fungerer som en del av virksomheten for øvrig, samtidig som det må tilrettelegges for at de som evaluerer evner å tre ut av denne nærheten med distanse og objektiv vurdering som ideal. Det er denne funksjonen som Vedung kaller återkoppling, og som jeg vil komme tilbake til i punkt 4.1.3.

Evaluering beskrives derfor godt via den didaktiske relasjonsmodellen (Bjørndal og Lieberg 1978), som en del av virksomheten som helhet, og skal behandles deretter. Det betyr at den i likhet med de andre kategoriene skal tilpasses og justeres i tråd med de andre, slik at evalueringen kan gi et riktig bilde av objektet og være med å tilrettelegge for at det som avdekkes kan gi læring. Dette innebærer at evalueringen derfor må planlegges i sammenheng med de andre kategoriene for å skape et best mulig grunnlag å evaluere ut fra. Arbeidet med kategoriene i sammenheng i denne fasen vil bidra til å kvalitetssikre alle ledd i en evaluering, og bidra til best mulig validitet og tilrettelegging for læring.

Qvotrup har også vist at det å legge til rette for læring er en komplisert sak, som innebærer å ta hensyn til ulike nivåer for læring, der graden av refleksjon er viktig. Evne til refleksjon over egen praksis er nært knyttet til evne til læring. Derfor er planlegging og tilrettelegging av evalueringen vesentlig for å sørge for at det blir lagt til rette for refleksjon på ulike nivåer i organisasjonen. Rom for refleksjon gjelder også forskerne selv. For å utøve profesjonalitet i evaluering kreves det også at forskernes eget arbeide utsettes for vurdering og refleksjon både av dem selv, og av miljøet rundt.

Selve grunnlaget for læring av en evaluering legges i forberedelsesfasen av evalueringsprosessen. Det er her tilretteleggingen av evalueringens bidrag til refleksjon og rasjonalitet i utførelse jamfør Dale skjer. Det som ikke tilrettelegges i planleggingsfasen kan ikke forventes å påvirke sluttproduktet av evalueringen. Det er i selve gjennomføringen at konsekvensene av god eller dårlig planlegging vises.

Den didaktiske evalueringsmodellen definerer forberedelsesfasen av en evaluering ved hjelp av de didaktiske kategoriene. Kategoriene viser både hva som må til av tilrettelegging innenfor hver fase for at det skal være mulig å lære av den, og hvilke forhold det er nødvendig å ta hensyn til for at evalueringen skal evne å evaluere det den er ment å evaluere. De didaktiske kategoriene viser også at de, ved å bli benyttet som utgangspunkt for planleggingen av en evaluering, bidrar til å bestemme evalueringens verdikriterier. Det er derfor særlig i denne fasen jeg mener at den didaktiske evalueringsmodellen viser sin eksistensberettigelse i evalueringssammenheng.

Vedung legger vekt på at forberedelsesfasen bør avsluttes av en formell beslutning før gjennomføringsfasen iverksettes og evalueringen for alvor settes ut i livet. Jeg deler dette synspunktet. Ved å foreta en beslutning tvinges forskere og oppdragsgivere til å

oppsummere og gjennomtenke planleggingen som er gjort enda en gang, og får gjennom det en forsikring om at de forberedelser som kunne gjøres er gjort, og at det materialet som er planlagt gjennomgått svarer på bestillingen som er gitt forskerne og evalueringen. Beslutning er derfor en fase i den didaktiske evalueringsmodellen som danner avslutningen på forberedelser og planlegging, og overgangen til gjennomføringsfasen.

#### **4.1.2 Gjennomføringsfasen**

I selve gjennomføringsfasen av en evalueringsprosess skal alle planlagte grep for å svare på evalueringens bestilling iverksettes. I tillegg til selve iverksettelsen, viser den didaktiske evalueringsmodellen at det i denne fasen skal utføres en revisjon av evalueringen, som er beskrevet utført på K2 nivået i henhold til Dales kompetansenivåer, og det forutsetter refleksjon som gir mulighet for omlæring i følge Qvortrups tredje ordens viten. Her skal det gjennomføres refleksjon over den planen for evaluering som er lagt i forberedelsesfasen, og den praksis som den har utløst.

Mange evalueringsoppdrag starter i gjennomføringsfasen. Hvis målet er å beskrive hva som skjer, gjennom å følge en virksomhet over tid, kan det virke uhensiktsmessig og vanskelig å legge planer på forhånd. Det kan også oppleves som manipulering og skapte føringer å ha synspunkter om hvordan virksomheten skal vurderes, og hva som skal velges av metoder og fokus, i forkant av en evaluering. Hvilken oppfatning de som utfører evalueringsarbeidet, har på forståelsen av hvordan evalueringer helst bør utføres, vil være avgjørende for hvilken vekt som blir lagt på forberedelsene.

Evalueringsrolle i organisasjonen vil bestemmes av hvilke forutsetninger den er gitt. Det som ikke er planlagt er det derfor vanskelig å ”hente inn” og tilrettelegge for i gjennomføringen. Den første fase i gjennomførelsesfasen innebærer å tilrettelegge for en rekke støtteprosesser som er nødvendige for å iverksette det som ble bestemt i forrige fase. Blant annet å sørge for at det finnes nok personalressurser til å utføre gjennomførelsesprosessen.

En viktig del av forberedelsesarbeidet til en evaluering er å sørge for at det er tilgjengelig nok ressurser til å utføre evalueringsoppdraget tilfredsstillende, slik at det sikres god kvalitet på sluttproduktet. Dette handler om profesjonalitet i utøvelse. En av mine foreløpige hypoteser i oppgaven var at kvaliteten i evalueringens utarbeidelse har betydning for hvilken

rolle den senere får i organisasjonen. Derfor er oppdragsgivere og ledelsens syn på organisasjonen utslagsgivende for hvilken rolle den får. Synet på organisasjonen får igjen følger for evalueringens organisering og hvordan den forberedes og tilrettelegges. For å kunne bedømme hvorvidt en evaluering er gjennomført for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”, er det nødvendig å se på hva utførelsen sier om hvilken organisasjonsforståelse som preger organisasjonen.

I følge Senge har aktørene og relasjonene/samhandlingen mellom dem innvirkning på organisasjonens evne til å lære. Disse kulturelle forutsetningene har derfor betydning for evalueringen. Involvering handler om tilrettelegging for deltakelse i alle deler av prosessen, slik den didaktiske relasjonsmodellen har vist. Denne involveringen er en nødvendig forutsetning for en vellykket gjennomføring.

Når utfall og resultat av gjennomføringsfasen foreligger, skal det vurderes om evalueringen har fungert etter sin hensikt så langt, og refleksjonen rundt dette skal så danne grunnlag for det Vedung kaller återkoppling. Jeg vil i det neste vise hva dette sier om veien videre i evalueringsprosessen.

### **4.1.3 Videreutvikling og nykobling**

Etter at gjennomføringen er gjort og resultatet foreligger, viser den didaktiske evalueringsmodellen at evalueringen som helhet ikke er gjennomført før en ny revisjon av evalueringsarbeidet er gjennomført. En av forutsetningene for at evalueringen utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”, er at evalueringen i seg selv utsettes for evaluering. Revisjonen som foregikk under gjennomføringsfasen er ikke nok. En av evalueringens viktigste funksjoner er å sørge for at det kan foregå refleksjon med distanse og utenfrablikk til evalueringsobjektet. Dette er en nødvendig forutsetning for at refleksjon over praksis skal skje, og at den refleksjon som finner sted skal være av en kvalitet som gjør at den kan føre til læring.

For å få til dette mener Vedung at det er nødvendig å legge opp til en fase i en evalueringsprosess der forskerne tar et skritt tilbake, ut av evalueringsarbeidet, og vurdere evalueringen utenfra. Vedung kaller den fasen for återkopplingsprosessen og jeg har omskrevet den til nykobling. I navnet ligger det at forskerne går ut av sitt eget arbeid og vurderer evalueringen i sin helhet, før den bearbeides og kobles på reformarbeidet eller



evalueringsobjektet på nytt igjen. Fasen innebærer derfor å gå inn og vurdere om det verktøyet som er utviklet og benyttet fungerer i følge sin intensjon og i henhold til de mål som er lagt for evalueringen. Dette innebærer å se hele evalueringsprosessen fra initiering og planlegging, til gjennomføring og utfall i sin helhet, for å se om evalueringen i gjennomføringsprosessen har fungert slik den var ment å fungere.

Bearbeidelse av evalueringen ligger også innenfor denne fasen, før den på nytt iverksettes og kobles på evalueringsobjektet. Det er derfor ikke samme evaluering som kobles på igjen, men en bearbeidet og videreutviklet utgave. Dette er en fase som det må forventes at forskerne selv tar tak i, samtidig som det sørges for at refleksjonen rundt denne fasen også kommer den enkelte medarbeider til del, jamfør behovet innenfor ”en lærende organisasjon”. Nykoblingsfasen forutsetter det Dale vil kalle rasjonalitet i eget arbeide. Å utsette seg for kritikk, og ha en kritisk holdning til egen forskning som gjør det mulig å endre praksis i henhold til nye erfaringer. Det er i dette arbeidet det største læringspotensialet ligger i evalueringer.

Nykoblingen forutsetter at det som må endres kan endres i den fasen av evalueringsprosessen man er i. Gjennomgangen av prosessen så langt har vist at det er forhold som ikke lar seg endre senere i prosessen, og som forutsetter god tilrettelegging i forberedelsesfasen. Likevel vil endringer i denne fasen i det store og hele bidra til å kvalitetssikre evalueringsprosessen som helhet og forbedre denne.

Nykoblingsprosessen forutsetter refleksjon på K3-nivået i følge Dales kompetansenivåer. Funksjonen er her å fremføre argumenter, å grunngi standpunkter, og å drøfte begreper som har med evalueringens utforming og gjennomføring å gjøre. Ved hjelp av tilbakemeldinger på det evalueringsarbeidet som er utført vil forskerne få viktige bidrag til dette arbeidet. I denne fasen forventes det at forskeren kan reflektere over forutsetningene for sine egne refleksjoner i den evaluering som er utført. Jeg har derfor koblet denne fasen med Qvortrups 3.ordens viten i den didaktiske evalueringsmodellen. Fungerte evalueringen i følge hensikten? Tilrettelegger den for læring slik at den kan fungere som et verktøy i ”en lærende organisasjon”? Gir den svar på det en var ute etter, og er det tilrettelagt for nok refleksjon hos den enkelte medarbeider? Eventuelt hvilke endringer bør gjøres før evalueringsprosessene fortsetter? Disse spørsmålene må besvares før nykoblingen iverksettes dersom denne delen av evalueringsprosessen skal tas på alvor.

I tillegg til nykoblingsprosessen anbefaler Vedung også å gjennomføre en prosess for metakontroll. Denne metakontrollen tar enda et skritt tilbake fra evalueringsobjektet og vurderer den nykoblingsprosessen som er utført, samt evalueringen i et videre perspektiv enn i det forrige. Jeg har koblet denne fasen i prosessen på Qvortrups 4.ordens vitensposisjon. Fjerde ordens viten betegnes som nevnt som ”metasystemisk viten”. Det handler om å forstå og reflektere over betingelsene for det systemet som skaper betingelser for dine refleksjoner over hva du kan. Evalueringenes forutsetninger, forankring, karakter og kvalitet i møte med omgivelsene er gjenstand for vurdering på dette nivået, sammen med spørsmål om hva som eventuelt bør endres ved den utførte evalueringsprosessen som helhet før neste evaluering.

Metakontrollen danner sluttfasen i hele evalueringen. Det er selvsagt mulig at også metakontrollen gjøres flere ganger i forbindelse med evalueringer av virksomheter som skjer kontinuerlig.

### *Oppsummering*

Modellen jeg har presentert i dette kapitlet er et resultat av arbeidet med denne oppgaven. Gjennom arbeidet med teoriene i kapittel 3 og særlig evalueringsteorien til Vedung, har jeg reflektert mye over at han har landet på ikke å definere noen verdikriterier for sin teori. Riktignok medgir han at evaluering bør innebære ett verdikriterium men han angir ikke hvilket.

Mitt engasjement i denne oppgaven har blant annet vært å definere innholdet i en god evaluering, og finne ut av hva som skiller en god evaluering fra en dårlig. Hvor går disse skillelinjene mellom god og dårlig, og hva sier dette om evalueringens utforming og gjennomførelse?

Jeg mener at jeg gjennom presentasjonen av den didaktiske evalueringsmodellen har nærmet meg en løsning på dette spørsmålet. Gjennom min didaktiske tilnærming, og gjennom utviklingen av den didaktiske evalueringsmodellen, mener jeg å ha definert kriterier og standarder som nettopp utfyller det jeg opplever som en mangel i Vedungs teori. De didaktiske kategoriene til Bjørndal og Lieberg (1978) vil, hvis de blir benyttet og evalueringen utviklet i tråd med intensjonene innenfor hver av dem, sørge for at evalueringen gir svar på de spørsmål som den er ment å svare på. Dette innebærer også at den vil måle/vurdere/evaluere ut fra de standarder og kriterier som forberedelsesfasen og planleggingen har lagt opp til.

Et hovedmoment i all evaluering er, som jeg har berørt tidligere, at den skal føre til læring i en eller annen form. Jeg har sett at læring i seg selv ikke holder som eneste verdikriterium i evaluering. Læring er ikke i seg selv nødvendigvis en ønsket prosess. Læring som fenomen kan derfor ikke stå alene som verdikriterium, men må følges av andre kriterier som bygger opp under og bidrar til å tilrettelegge for at den læring som skjer er av en art som er ønsket. Disse kriteriene mener jeg å ha funnet gjennom de didaktiske kategoriene. I tillegg til å hevde at jeg gjennom Vedungs definisjon av evaluering og Dale og Qvortrups syn på læring har vist at læring er et nødvendig verdikriterium i all evaluering, mener jeg i tillegg å ha vist at de didaktiske kategoriene utfyller og støtter opp om læring som kriterium og gir standarder å vurdere og måle ut fra.

Jeg mener derfor at dette kapitlet har vist hvordan en evaluering ideelt sett bør utarbeides og gjennomføres for å sikre kvalitet og derigjennom sørge for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”. Den didaktiske evalueringsmodellen, via en didaktisk innfallsvinkel til evalueringer, gir fokus på planlegging og forberedelse som igjen kan sikre kvalitet i utarbeidelse, gjennomføring og resultat av evalueringer. Modellens mange faser viser også hvordan evalueringen bør tilrettelegges for å sikre at den skal føre til læring hos den enkelte aktør og derigjennom ”den lærende organisasjonen”. For å sikre at tilretteleggingen for læring er best mulig overfor den enkelte medarbeider, er det en forutsetning at det blir utviklet et program for hvordan evalueringen er tenkt gjennomført. Den didaktiske evalueringsmodellen er ment å fungere som et redskap i arbeidet med å utvikle et program.

I kapittel 5 vil jeg bruke den didaktiske evalueringsmodellen som verktøy for analysen av evaluering av reform av trosopplæring. Jeg vil vurdere testevalueringen min i forhold til modellen, og til slutt se på evalueringen som helhet og hva den kan si oss om det lærende utbyttet som er beskrevet av forskerne.

## 5. Analyse

I dette kapitlet vil jeg drøfte hva teorien i kapittel 3 og 4 sier om evalueringsoppdraget i reform av trosopplæring, og måten det er utført på. Dette vil jeg gjøre ved hjelp av den didaktiske evalueringsmodellen som jeg presenterte i forrige kapittel. Ved å gå gjennom de ulike fasene av modellen vil jeg belyse hva teorien kan si oss om evalueringen innenfor hver fase, og ved å gå gjennom hele modellen vil jeg drøfte hvordan evalueringen i sin helhet er utført. Dette kapitlet vil derfor være en direkte anvendelse av den didaktiske evalueringsmodellen sett i lys av problemstillingen:

**Hvordan bør en evaluering utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”?**

Som avslutning på kapitlet vil jeg oppsummere hva konklusjonen på drøftingen gir av informasjon om betydningen av en prospektiv innfallsvinkel til evalueringsfeltet, og hva den didaktiske evalueringsmodellen kan bidra med i arbeidet med å tilrettelegge evalueringen slik at den sørger for best mulig grunnlag for læring i en ”lærende organisasjon”. Gjennom konklusjonen på dette vil jeg så i kapittel 6 oppsummere oppgaven, og begrunne hvorfor jeg mener at læring er et nødvendig kriterium i all evaluering.

### 5.1 Anvendelse av den didaktiske evalueringsmodellen

Presentasjonen av den didaktiske evalueringsmodellen i forrige kapittel har vist at en evaluering består av mange ulike faser, der behandlingen av de enkelte fasene har betydning for kvaliteten på evalueringen i sin helhet. Gjennomgangen av modellen har også vist at forberedelses- og planleggingsfasen legger grunnlaget for gjennomføring og sluttprodukt. Jeg vil nå gå gjennom de ulike fasene i den didaktiske evalueringsmodellen, slik jeg gjorde i kapittel 4, men nå for å analysere hvordan evaluering av reform av trosopplæring har forholdt seg til de ulike fasene, slik jeg har kunnet lese det ut i fra forskernes egne vurderinger i rapportene. Jeg forutsetter at leser har satt seg inn i hva jeg legger i de ulike fasene, slik jeg har beskrevet dem i kapittel 4.

### 5.1.1 Forberedelsesfasen

Jeg vil i det følgende beskrive forhold som har betydning for utarbeidelse av evalueringen, og testevalueringens forberedelsesfase. Forberedelsesfasen innebærer som nevnt initiering, planlegging og beslutning, som er de tre første fasene i den didaktiske evalueringsmodellen. Initiering vil i hovedsak dreie seg om forhold knyttet til formål og intensjon ved evalueringen fra oppdragsgivernes side, og hvordan planen må tilrettelegges for å ivareta dette. Under ”planlegging” skal alle forhåndsbestemmelser i startfasen operasjonaliseres.

#### *Initiering*

I reform av trosopplæring er formålet uttrykt gjennom styringsdokumentene som følger reformen. Hva disse sier er derfor vesentlig for all videre planlegging, og hva slags tilrettelegging for læring som kreves for å virkeliggjøre formålet. Dette planleggingsarbeidet forutsetter det nivået for læring som Qvortrup beskriver som fjerde ordens viten. Forhold knyttet til programmets rolle og innflytelse i og utenfor organisasjonen, samt visjoner for organisering og læring i organisasjonen og hva det innebærer for en ”lærende organisasjon”, hører inn under refleksjonen som bør foregå i denne fasen.

I reform av trosopplæring benyttes evalueringen som et styringsredskap, for å sørge for at reformledelsen benytter forsøksfasen i tråd med reformens intensjon, som var å sikre alle menighetene i Den norske kirke en mulighet til å tilby en systematisk trosopplæring til alle døpte, og for å sikre at Stortinget ble oppdatert på reformen gjennom forsøks tiden. Oppdragsgiverne legger i denne fasen alle føringer for økonomi, vilkår, og mål for evalueringen, og gir gjennom dette rammene og forutsetningene for den evalueringsprosess som er tenkt gjennomført. I evalueringen av reform av trosopplæring er oppdragsgiver Stortinget ved Kirkerådet. Stortinget initierte reformen og med den også evalueringen, men la som nevnt få føringer for hvordan evalueringsoppdraget skulle utføres.

Den eneste føringen som ble lagt i vedtaket fra Stortingsbehandlingen når det gjaldt evalueringen var at reformen skulle følges av en forskningsbasert evaluering. Det ble ikke fra Stortingets side sagt noe om forventet læringsutbytte, eller hva man ønsket å oppnå med evalueringen. Hva som ligger i betegnelsen forskningsbasert har det vært opp til reformledelsen og evalueringsforskerne å bestemme. De har tolket oppdraget slik at det skulle være en følgeforskning som skulle gi Stortinget informasjon om reformen underveis i forsøks tiden, og som skulle ende opp i en summativ rapportering mot slutten av

forsøks tiden. Ideelt sett burde oppdragsgiverne som initierte reformen også sørget for tydeligere føringer på hva formålet med følgeforskningen skulle være, og hva det forventede resultatet av den skulle være.

Siden forutsetningene fra Stortingets side ikke gir sterkere føringer om valg av evalueringsmetoder og hva evalueringen skal vise til av sluttprodukt, har reformledelsen hatt stor frihet til å tilrettelegge evalueringen. De fikk av Kirkerådet mandat til å iverksette evalueringen og engasjere evalueringsforskere. Dette innebærer i realiteten at Stortinget i stor grad har lagt opp til en selv-evaluering. I følge Vedung kan selv-evalueringer være nyttige virkemidler i ”lærende organisasjoner” som er opptatt av at alle medlemmer og aktører skal være lærende. Evaluering innebærer nettopp selvrefleksjon, det selvrefererende læringsgrunnlaget. Organisasjonens rasjonelle fornyelse forutsetter denne evnen til å se seg selv, til å ha seg selv som kollektivt tema (Dale 1998).

Vedung argumenterer imidlertid for at eksterne forskere i tillegg kan bidra med et verdifullt utenfrablikk. Maktkonstellasjoner og hegemoni kan hindre evalueringsforskere i å utføre oppdraget i henhold til formålet. Med tanke på den erfaring Stortinget har fra evalueringene av reform 94 og L97 i skolesektoren, er det overraskende at det politiske miljøet ikke har vært mer opptatt av å sikre at evalueringen av reform av trosopplæring ble foretatt av forskere som stod utenfor reformorganisasjonen. For å skape best mulig grunnlag for objektivitet og utenfraperspektiver i denne type evalueringer ville det vært ønskelig at Stortinget selv hadde opprettet forskningsmiljøet, og lagt tydelige føringer for hva de ønsket at evalueringen skulle gi svar på. De overlot i stedet oppgaven til Kirkerådet, som igjen gav den videre til Styringsgruppen. Heller ikke Kirkerådet har vært opptatt av å sikre sin egen og prosjektledelsens uavhengighet til forskerne og derigjennom forskernes objektivitet og kvalitet. Ved å overlate utnevning og valg av forskere til prosjektledelsen kan de kritiseres for en viss ansvarsfraskrivelse.

I realiteten har dette medført at de som har sittet nærmest ledelsen av forsøksfasens utvikling og gjennomføring, både har opprettet evalueringsgruppen og lagt føringene for hvilken rolle de skulle ha. Styringsgruppens sekretariat har dermed hatt sterk kontroll med valg av forskere og hvilke rammer de skulle arbeide under. Med tanke på hvilket nybrottsarbeid vi her snakker om, og omfanget på reformen både når det gjelder antall involverte og hvilke økonomiske rammer som er tildelt, er det gitt dem et stort ansvar. Sekretariatet må i prosessen med å utvikle og ”bygge” forsøksfasen forventes å ha utviklet en sterk eierfølelse

til reformarbeidet. Likevel er de her satt til å ha oppgaven med å lønne og administrere de som er satt til å vurdere deres eget arbeid.

Når det er sagt så har forskerne likevel vist at de kan stille seg kritiske til sekretariatet og reformledelsens arbeid. Selv om de må omtales som kritiske venner, fordi de relativt sett har presentert kritikk ”pent” innpakket i hyggelige perspektiver, har de likevel vært til dels sterkt kritiske til de strategier som er lagt og de prosesser som er gjennomført. Dette viser at utenfrablikk er forsøkt ivaretatt og at forskerne har sett det som sin oppgave å bidra til å utvikle reformen. Dette innebærer igjen at de har lagt vekt på å skape muligheter for at organisasjonen skulle lære av de erfaringene som er gjort.

Vedung argumenterer for at spørsmålet om hvem som skal ha evalueringsoppdraget må vurderes fra evaluering til evaluering, og at forskere med nærhet til organisasjonen bør velges der det er gunstig med godt kjennskap og innenfraperspektiver hos forskerne. Det finnes gode argumenter for at forskerne i denne evalueringen bør stå nært prosjektorganisasjonen, blant annet fordi oppdraget er spesielt med tanke på innhold og mål. Likevel er det overraskende, særlig med tanke på Stortingets opptatthet av å bevare en bred folkekirke, og at reformen skal ha fokus på å nå bredt ut til befolkningen, at de velger å la prosjektorganisasjonen selv styre hvem som skal vurdere forsøksfasen og hvordan det skal gjennomføres.

### **Forskernes rammer**

De rammer som er gitt forskerne i evalueringens mandat gir igjen rammer for evalueringens organisering og omfang, og har betydning for hvorvidt forskerne kan bidra til å virkeliggjøre formålet med evalueringen. Forskernes rammer har betydning for kvaliteten på evalueringen, og det er derfor vesentlig å kjenne til i utviklingen av evalueringsprogrammet. Rammene de må arbeide under er også i stor grad bestemmende for hvilken rolle evalueringen får i organisasjonen og overfor den enkelte aktør. De har derfor også innflytelse på hvilket grunnlag for læring den gir i utarbeidelse og gjennomføring i en ”lærende organisasjon”.

De rammene forskerne i reform av trosopplæring har hatt å forholde seg til er lagt av styringsgruppen ved sekretariatet, og er som nevnt nedtegnet i avtalen mellom styringsgruppen og dem. Det går ikke fram av avtalen at det er forskernes oppgave å utvikle evalueringsskjemaer eller rapporteringsskjemaer. Rapporteringene som sekretariatet har utarbeidet har vært benyttet i forskernes vurderinger, og har derfor dannet grunnlag for

kvaliteten på evalueringen. Disse skjemaene eller evalueringene skal derfor svare på bestillingen som er gitt forskerne. Siden det er gitt så åpne føringer fra Stortinget, er det bestillingen som er gitt av Kirkerådet og Styringsgruppen forskerne først og fremst skal forholde seg til. Denne er så langt jeg kan se innfridd. Det må derfor kunne sies at forskerne har opptrådt innenfor de rammer som er gitt dem. Etter mitt syn må de vel heller kritiseres for ikke å ha utnyttet rollen sin mer enn de har, og dermed gitt seg selv bedre mulighet til å utøve profesjonalitet i arbeidet. Dette mener jeg viser seg i følgende forhold:

Hvilket materiale som skal danne grunnlag for evalueringen, og hvordan det skal utarbeides, henger nært sammen med hva evalueringen er ment å vise til. Hva skal måles og for hva? Og hva er evalueringen i seg selv ment å brukes til? Forventninger oppdragsgiverne har til forsøksfasen og hva de uttrykker at de vil at evalueringen skal gi informasjon om, er derfor viktig å finne ut av. Bruken av tall på oppslutning til tiltakene i startfasen av forsøkstiden, viser at dette spørsmålet er det tatt hensyn til av reformledelsen. Det har vært viktig å vise Stortinget foran hver budsjettbehandling at forsøkstiden så langt har vært vellykket. Dette ble blant annet gjort med bakgrunn i antall forsøksmenigheter og oppslutning om tiltakene. Forskerne har senere vist til at tallene er misvisende fordi rapporteringsskjemaene er tolket feil. Hvis tall på oppslutning var viktig for å gi tilbakemelding om forsøksfasen til de politiske oppdragsgiverne, burde ikke da forskerne ha vært mer opptatt av validiteten i dette materialet?

Dette burde vært sett på som viktig for å besvare Stortingets intensjon med å bestille en forskningsbasert følgeforskning, og den delen av avtalen med styringsgruppen som sier at den skal beskrive trosopplæringens utforming, utvikling, og konsekvenser. Det gjelder også det pretest-materialet som ble utviklet i starten av forsøksfasen, og som viste seg ikke å holde god nok kvalitet til å kunne brukes. For å sikre et grunnlag å vurdere forsøkstiden opp mot burde det vært gjort ytterligere forsøk på å innhente de opplysninger det var behov for fra starten av. De slutninger som trekkes på vegne av det materialet som er benyttet er i stor grad preget av syning og grunne påstander.

Forskerne har ved flere rapporteringer gitt tilkjenne at det datamaterialet de har hatt til rådighet ikke har gitt validitet i svar blant annet fordi det har vært uklarheter i tolkningen av hvordan rapporteringsskjemaene skulle fylles ut. Forskerne har pekt på at prosjektledelsen har hatt utfordringer i utviklingen av rapporteringsskjemaene, at de har hatt for lite ressurser og derfor vært svært presset i tid, og at sekretariatet til tross for disse forhold likevel har



utviklet et datamateriale som er tilfredsstillende. Ved gjentatte rapporteringer har de gitt sekretariatet ros for utbedringer av rapporteringsgrunnlaget. Dette kan sees som et tegn på at forskerne ikke har tatt rollen som kvalitetssikrer av sitt eget grunnlagsmateriale, men overlatt oppgaven til sekretariatet.

Forskernes mulighet for å gå inn å kvalitetssikre materialet på forhånd har jeg ikke sett nærmere på. Muligheten for at de kan være hindret av prosjektledelsen i forhold til å utvikle rapporteringsmaterialet selv, er til stede. De økonomiske rammene de er gitt kan for eksempel ha skapt begrensninger. Det går som sagt heller ikke fram av avtalen at det er en del av forskernes oppdrag å utvikle rapporteringsgrunnlaget. Likevel mener jeg at en så viktig del av et evalueringsarbeid, som å sikre at alle data kan gi svar på bestillingen av evalueringen, bør utvikles og kvalitetssikres av forskerne selv. Dette arbeidet bør utføres helt i begynnelsen av en evaluering, før innhenting av datamaterialet starter. Dette bør, slik jeg ser det, være en naturlig del av forberedelsen i alle evalueringsarbeider.

### **Evaluerings mål**

Som start på en evalueringsprosess hører det også hjemme å bestemme mål for evalueringen. Formålet med evalueringen skal operasjonaliseres i mål som skal danne grunnlag for en felles grunn å arbeide ut fra for alle medlemmene i organisasjonen.

Hvilke muligheter en har for å hevde å ha innfridd det målet som er lagt for reform av trosopplæring er ikke lett å svare på. Målet var som nevnt: *å utvikle en systematisk trosopplæring som fremmer kristen tro, gir kjennskap til den treenige Gud og gir hjelp til livstolkning og livsmestring for alle døpte i alderen 0-18 år, uansett funksjonsnivå.* Å operasjonalisere et slikt mål, for å kunne finne ut om det er nådd, anser jeg å være en nesten umulig oppgave. Så har da forskerne heller ikke benyttet reformens mål for sitt evalueringsarbeid. Som vi har sett er Vedung også kritisk til å benytte mål som utgangspunkt for evalueringer, fordi han mener at forskerne lett lar seg lede mot målet, og dette hindrer etter hans mening muligheter for å se blant annet bieffekter i en evaluering. Forskerne har i stedet for å oppstille mål for sin evaluering, sammenfattet sitt mandat i et overordnet spørsmål: *Har forsøksfasen vært vellykket, og er intensjonen oppfylt?* For å svare på det har de laget en modell for å måle endring, og ved hjelp av tre hypoteser for endringer har de undersøkt forsøksfasen for å avdekke om funnene kan si noe om reformen har ført til verifisering eller falsifisering av hypotesene.

Det er på mange måter forståelig at forskerne har unngått målstyringen, men det er grunn til å stille spørsmål ved om hypotesene gir et mer valid bilde av forsøksfasen. Kan hypotesene om endring sies å svare på det overordnede spørsmålet? Og kan operasjonaliseringen av hypotesene for endring gi de svar de er ment å gi? Hypoteseformuleringene kan virke lite reflekterte når det gjelder realiserbarhet. Å besvare om en forsøksfase har vært vellykket stiller store krav til forberedelse og operasjonalisering. Hva sier dessuten forskernes måte å løse evalueringsoppdraget på om evalueringens plass i organisasjonen og om didaktisk rasjonalitet i utførelse?

Dale har vist at å gjennomføre opplæring med didaktisk rasjonalitet innebærer målrettet læring, og gjennom det felles referanseramme og forståelseshorisont. Behovet for å samles om felles mål er derfor viktig innenfor all virksomhet som har læring som mål, og spesielt i ”lærende organisasjoner” som har som mål å legge til rette for læring hos den enkelte medlem/aktør. Det er derfor også et viktig anliggende for evalueringer. For å ivareta didaktisk rasjonalitet i utførelse, har Dale påpekt betydningen av konsistens. Dette innebærer også at eksterne forskere som kommer utenfra må forholde seg til de samme mål som organisasjonen styrer etter, hvis evalueringen skal fungere rasjonelt og fungere overfor medlemmene av organisasjonen. En evaluering kan få betydning som en medierende artefakt (Säljö 2006) overfor den enkelte profesjonelle aktør dersom den blir en del av deres felles praksis, og evaluerer denne i lys av felles mål. På denne måten vil evalueringens læringseffekt ivaretas best mulig. Den vil også på denne måten få betydning for bygging av kultur i organisasjonen som helhet, fordi den vil bli et felles eie og et felles verktøy for læring og utvikling.

Initieringen har følger for evalueringens rolle i organisasjonen. Alle forutsetninger og rammer som er nevnt så langt i dette kapitlet legger grunnlaget for den læring som evalueringen er ment å gi. Det er grunn til å spørre om Stortingets intensjon med evalueringen er fulgt opp gjennom den bestillingen forskerne har fra reformledelsen, og om det datamateriale som danner grunnlag for evalueringen er i stand til å oppfylle denne intensjonen. Svikt i denne fasen av et evalueringsarbeid gir et lite heldig utgangspunkt for det videre arbeidet i evalueringen.

## **Planlegging**

I arbeidet med å utvikle et program for evalueringen vil det som nevnt være naturlig, slik den didaktiske evalueringsmodellen har vist, å ta utgangspunkt i de didaktiske kategoriene: Mål, Målgruppe, Innhold, Rammer/kontekst, Arbeidsmåter, Gjennomføring, og Evaluering. Kategoriene bidrar til å definere kriterier i evalueringen ved å angi de standarder evalueringen skal måles etter. Hvordan planleggingen av evaluering av reform av trosopplæring er utført i henhold til dem vil jeg omtale i det følgende. Jeg vil derfor gå inn i de ulike kategoriene og beskrive hva min analyse om evalueringen sier om innholdet i dem.

## **Operasjonalisering av Mål**

I denne fasen skal det utledes mål av den overordende målformuleringen. Denne prosessen handler om å formulere mål som kan operasjonaliseres og derigjennom besvares gjennom evalueringen. Det er her noen av de største utfordringene i evalueringen av reform av trosopplæring ligger slik jeg ser det. Operasjonaliseringen av mandatet som forskerne er tildelt er sammenfattet i et overordnet spørsmål:

### ***Har forsøksfasen vært vellykket, og er intensjonen oppfylt?***

Med dette som utgangspunkt har forskerne som nevnt valgt å undersøke hvilke endringer forsøksfasen har medført for Den norske kirkes trosopplæring. Med endringer har forskerne forstått både kvalitative og kvantitative virkninger som inntreffer, dels som følge av reformens nasjonale virkemidler, dels av lokale forhold, og dels som en følge av andre sider ved samfunnsutviklingen. Forskerne mener å beskrive og forklare hva endringene har bestått i, samt hvordan og hvorfor de skjedde. Utgangspunktet for trosopplæringsreformen er forstått av forskerne som en serie intensjoner om å endre etablert praksis. Den metodiske utfordringen i evalueringen har vært både å identifisere endringsimpulser, og vise hva slags endringer de har skapt. Forskerne har laget en modell for å analysere endringer i prosjektorganisasjonen. Modellen angir tre hypoteser om endring:

*I forsøksfasen har forsøksmenighetene vært preget av en læringsprosess, der systematisk utprøving av ulike typer tiltak har preget virksomheten.*

*I forsøksfasen har forsøksmenighetene tilpasset seg eksterne krav og forventninger gjennom ulike likedannende prosesser.*

*Med forsøksfasen som ramme har forsøksmenighetene så langt mulig søkt å videreføre egne tradisjoner for tros- og dåpsopplæring.*

Hypotesene som er presentert er alle innrettet mot å vise hva slags endringsdynamikker som gjør seg gjeldende.

Konkret har forskerne stilt følgende spørsmål: Hvilke tiltak ble satt i gang, hvilke målsettinger lå bak disse, og hvordan ble forsøkene organisert og dokumenterte? Hvilke erfaringer og resultater har forsøkene gitt? Hvilket innhold hadde de tiltakene som ble iverksatt? Hvordan var oppslutningen, hvor bredt nådde man ut med trosopplæringstilbudene? I hvilken grad og på hvilken måte ble trosopplæringen tilpasset lokale variasjoner i lokal kultur og kirkelig struktur? Hvordan har reformens støttetiltak, som mentorordning m.m. virket? Hvordan skal aggregerte erfaringer fra forsøksfasen kunne legges til grunn for reformens permanente gjennomføring?

Det første spørsmålet er kanskje det som er best besvart i forsøksstiden, og det som har best validitet i svar. Spørsmålet er operasjonalisert ved en rekke rapporteringer fra prosjektledere om det enkelte prosjekt. Dette er godt dokumentert, og jeg er her enig med forskerne i at det er gjort et godt grunnlagsarbeid i dette materialet. Men hva vet man egentlig om hvorvidt forsøksfasen har vært vellykket ved å besvare dette? Materialet gir godt kjennskap til prosjektenes innhold, organisering og ledelse, men siden vi ikke kjenner menighetens situasjon før forsøksfasen ble iverksatt, kjenner vi heller ikke til hva i prosjektene som skyldes forsøksfasen og hva som kunne vært gjennomført uten reformen. Når det også oppgis av forskerne å være en stor andel av prosjektene som oppgir å ha videreført tidligere påbegynte tiltak, som ikke var nye fra starttidspunktet, underbygger dette usikkerheten om nettopp dette. Det må likevel sies å være en viktig kartlegging i en forsøksfase som her er gjennomført.

Hvilke erfaringer og resultater forsøkene har gitt, og hvordan oppslutningen har vært, er etter forskernes egne vurderinger ikke godt nok operasjonaliserte. Misforståelse og forskjellige tolkninger preger de innrapporterte svarene. Dette tyder enten på dårlig kommunikasjon av datamaterialet, eller dårlig utarbeidede skjemaer. Kontroll spørsmål og andre kvalitetssikringssystemer i tillaging av datamaterialet kan her ikke være godt nok benyttet. Siden dette er svært viktig informasjon å hente ut av forsøksfasen med tanke på å gi råd om hvordan aggregerte erfaringer skal legges til grunn for reformens permanente

---

gjennomføring, er det synd og overraskende at det ikke er lagt mer vekt på å kvalitetssikre denne delen av reformarbeidet.

De spørsmålsstillingene som forskerne her har stilt og de operasjonaliseringer som er gjort, er derfor etter mitt syn ikke tydelige når det gjelder hvorvidt de funn de gjør kan måle om forsøksfasen har vært vellykket. Men det er kanskje ikke så rart, for hvordan kan man egentlig måle det? Hva skal være kriteriene for en vellykket forsøksfase? Forskerne har her prøvd å unngå å svare direkte på dette, og også på den målsetting som er lagt for hele reformen - med alle de komplikasjoner det har når det gjelder mangel på evaluering i forhold til det mål reformen styres etter, og felles mål og referanseramme for samtlige aktører.

### **Målgruppe**

I reform av trosopplæring er en ”lærende organisasjon” benyttet som organisasjonsteoretisk plattform. Når denne reformen skal evalueres ligger det da nært å hevde at målgruppen for evalueringen bør være alle aktører og medlemmer i organisasjonen.

Rapportene har vært tilgjengelige for alle ved at de har vært offentliggjort på nettet. Likevel henvender forskernes rapporter seg når de velger språk og formuleringer til de som har kjennskap til kirken som organisasjon, og aller mest til oppdragsgivere og reformledelsen. Utviklingen av de nettbaserte løsningene, har etter min vurdering vært av det beste ved reformarbeidet. Likevel har ikke evalueringen verken i sin organisering eller presentasjon henvendt seg til alle medlemmer og aktører i organisasjonen. Vet at forskerne i rapportene først og fremst har hatt fokus på brukere og tilsatte i prosjektorganisasjonen, har de mistet viktig informasjon ved ikke å legge opp til innspill fra aktører fra andre deler av organisasjonen. Evalueringen har derfor heller ikke fått betydning som samlende felles referanseramme og refleksjonsgrunnlag. Det nærmeste forskerne kommer disse anliggende er evalueringen av de menighetene som ikke har vært forsøksmenigheter. Grunnlaget her er likevel for lite til å gjøre opp for manglene.

Det er grunn til å spørre om dette er en heldig strategi med tanke på at reformtiden var ment å føre til implementering av resultatet av forsøks tiden i hele organisasjonen, og med tanke på at et av oppdragene til forskerne var å gi råd om overgang fra forsøk til reform.

Forskerne hevder at reformen hadde gode vilkår i organisasjonen, siden tiden i følge dem var moden for reformen. Likevel skapte den også utfordringer, da den brakte mye ressurser til få

og dermed opplevdes urettferdig i deler av organisasjonen. Også det faktum at Stortinget i noen grad forskjøv maktbalansen i organisasjonen ved å tildele økonomiske ressurser til prosjektledelsen, og med det gikk utenom den opparbeidede linjeledelsen, skapte utfordringer som forskerne burde gitt fokus med tanke på den senere implementeringen. Stortingets ”bottum up”- organisering skapte reformen som ”en stat i staten”. Dette må forventes å få betydning for implementeringen.

Hvordan forskerne bedre kunne bidratt til å gi et mer helhetlig bilde av forsøkstiden vil jeg komme tilbake til under arbeidsmåter. Det handler om valg av metode og å bestemme hvem det skal hentes informasjon fra i evalueringen. Valg av målgruppe handler derfor om hvem rapporteringene og evalueringen utarbeides for, og derigjennom hvem som forventes å skulle lære av den.

Reformen har alle døpte fra 0-18 år som målgruppe. Det er sagt lite om denne målgruppen fra et pedagogisk ståsted i evalueringsrapportene. I lys av det som har skjedd i skolen, og som er den opprinnelige grunnen til reformen i utgangspunktet, er det flerkulturelle miljøet som barnas oppvekstmiljøer representerer en viktig faktor å ta hensyn til for reformens oppdragsgivere og ledere. Skolen har i dag som mål å utvikle kritisk reflekterende og demokratiske individer<sup>3</sup>. Etter å ha lest utkast til ny plan for trosopplæring i Den norske kirke, som er utarbeidet på grunnlag av erfaringer fra forsøkstiden, har jeg lett etter formuleringer som ivaretar disse perspektivene overfor målgruppen. Begreper som tvil og kritisk holdning/refleksjon er ikke med i utkast til ny plan. I forhold til disse perspektivene mener jeg at forskerne, ved hjelp av fokus på en bredere målgruppe, kunne fanget opp og hjulpet reformledelsen til å få et bredere kjennskap til målgruppen i reformen. Dette handler i tillegg om å sikre at elevene har en felles kontekstualitet rundt livet som helhet, som også må oppfattes som en viktig del av å utvikle livskompetanse.

Å definere målgruppe i utviklingen av et program for evaluering har betydning for språk og formidling som må tilpasses målgruppen. Fordi medarbeiderne i Kirken, som er ment å lære av denne evalueringen, har forskjellig kompetanse, har organisasjonen utfordringer med å tilrettelegge slik at den kunnskapsproduksjon som evalueringen representerer faktisk blir en

---

<sup>3</sup> ”...å inngi respekt for kjensgjerninger og saklig argumentasjon - og oppøve kritisk sans til å gå løs på gjengs tenking, innarbeidede forestillinger og bestående ordninger” (utdrag fra L06 Generell del).

del av medarbeidernes kompetanse og derigjennom kunnskapsproduksjon for hele organisasjonen. Læringssystem med didaktisk rasjonalitet er avhengig av at medlemmene selv – inklusive ledelsen – bruker teori, tenker og kommuniserer i begreper i forhold til opplæringens funksjonshandlinger og eventuelt mangel på samme (Dale 1998).

## **Innhold**

I evaluering av reform av trosopplæring er det forsøksfasen som skal vurderes, og om den gjennomføres i tråd med intensjoner og mål. Men hva er egentlig innholdet i evalueringen av reformen? Organiseringen av forsøksfasen? Oppslutning? Systematisering av trosopplæring? Trosopplæringens innhold? Det er litt vanskelig å se at dette er avklart og avgrenset i evalueringsarbeidet.

Forskerne har løst dette ved å peke på alt i forsøkstiden som peker i retning av at det er skjedd en endring. Utfordringen deres har også vært at innholdet heller ikke har vært avgrenset i reformen. Forsøksfasen har vært preget av å la ”de tusen blomster blomstre” i følge dem selv, og i liten grad av å gå inn og påvirke/manipulere og skape retning i utprøvingene. Det er derfor et paradoks at de selv beskriver forsøksfasen som et eksperimentelt design, da de slik jeg ser det nettopp fremhever at det ikke er det. Her foregår det ingen manipulering og styring eller testing, kun beskrivelser og rapporteringer av hva som skjer.

Forskerne spurte etter mer fokus på innhold i første fase av forsøkstiden, fordi de mente at det var vanskelig å få øye på forskjellen i trosopplæring og annet barne- og ungdomsarbeid. Dette ble det i sterkere grad lagt vekt på å se etter senere i forsøksfasen, både i utvelgelse av prosjekter og i rapporteringsskjemaene. Dette ble ikke minst etterspurt som utgangspunkt for arbeidet med ny plan for trosopplæring. Det ble også av forskerne etterlyst et større fokus på systematisering av de erfaringene som forsøksfasen gav. Jeg har spurt meg om ikke forskerne burde ha bidratt til denne systematiseringen gjennom sitt arbeid. Kunne fokus på endringer vært tonet ned til fordel for en sterkere konsentrasjon på å belyse innholdet?

I stedet for å bidra til systematiseringen har forskerne hevdet at fagmiljøene burde vært dratt inn i systematiserings- og utviklingsarbeidet med tanke på en ny nasjonal plan for trosopplæring. Etter min oppfatning burde en del av evalueringens grunnlagsmateriale vært utarbeidet med tanke på å belyse innholdet i reformen. Det ville gitt større tyngde til hele evalueringen. Slik den framstår nå, så er evalueringen beskrivende og mest orientert om et

ytre perspektiv innad i prosjektorganisasjonen. Hva som skal læres gjennom evalueringen er ikke definert på forhånd, men er tenkt utviklet gjennom forsøksfasen i likhet med innholdet i reformen forøvrig.

Forsøksfasen har vist at prosjekter har oppgitt å ha benyttet dåpsopplæringen av 1991 til hjelp for å bestemme innholdet i prosjektene i mangel av noe annet. Selv om det er sannsynlig at forrige plan er benyttet mye, så har det ikke vært formulert som en felles sannhet. Evalueringen har spurt etter innholdsbestemmelse, men ikke medvirket aktivt til å løse innholdsspørsmålet. Innholdet henger nøye sammen med målet, og som nevnt under mål er det i følge Dale en forutsetning for rasjonalitet i utførelse at medlemmer og ledelse av en organisasjon har den samme forståelseshorisonten. Avklaring av innholdet som evalueringen skal orienteres om er i følge den didaktiske evalueringsmodellen nødvendig for å skape felles forståelse hos medlemmene i organisasjonen. Refleksjonen (K2) hos den enkelte medarbeider, og mellom kollegaer i organisasjonen i selve gjennomføringsfasen, bør som nevnt ikke handle om hva slags innhold som skal evalueres, men heller hvordan innholdet presenteres og behandles i organisasjonen.

### **Rammer/kontekst**

De økonomiske rammene som oppdragsgiverne tildeler evalueringsformålet vil begrense eller gi rom for kvalitet i alle deler av prosessen. Jeg har tidligere nevnt at forskernes oppgave her blir å definere behov i forhold til oppdraget, og gi tilkjenne de rom eller begrensninger som rammene gir. Dette er nødvendig siden kostnader knyttet til evalueringer er avhengig av oppdragsgivernes intensjon og behov vedrørende bruk, slik at det vil være nødvendig med tilpassninger i forhold til hver enkelt evaluering. Det må derfor innebære at forskere også må anse det som sin oppgave å påvirke og informere om hvilke muligheter eller konsekvenser en kostnadsramme vil ha.

Forskerne er selv tildelt mer midler for hvert år gjennom hele forsøksperioden. Om dette har vært gode nok økonomiske rammebetingelser sies det ingenting om. Det bør derfor kunne forutsettes at det har vært en kostnadsramme som har vært tilfredsstillende for å utføre oppgavene i tråd med deres egen opplevelse av oppdraget. Likevel kan vi ikke være sikre på at det medfører riktighet. Forskerne gir lite eller ingen melding i sine rapporter om egne rammebetingelser, og jeg har ikke gjort undersøkelser på det. Den økonomiske rammen for hvert år er imidlertid oppgitt, og utgjør 3,5% av hele reformbudsjettet i forsøksfasen.



Hvorvidt det er en liten eller høy andel av ressurser synes jeg det er vanskelig å vurdere. Siden mange av de kritiske bemerkningene jeg har til evalueringen handler om kvalitetssikring og utforming av datamaterialet som har vært hentet inn, mener jeg forbedringene jeg nevner i liten grad handler om økte økonomiske ressurser. Tid til utarbeidelse vil selvsagt koste, men slik jeg ser det kunne mye av tiden som for eksempel har vært benyttet til å analysere endringer, med fordel vært benyttet på å utbedre dette. Kanskje handler det derfor mer om valg og organisering enn om økonomiske rammer.

Det ble tidlig rapportert om utfordringer med verktøyet som ble benyttet til innhenting av data. Verktøyet ble imidlertid stadig utbedret og justert i henhold til behovet, og på slutten ble det omtalt som tilfredsstillende i forhold til formålet av forskerne selv. Det er derfor grunn til å tro at forskerne i liten grad kan ha sett på disse utfordringene som begrensninger i sine rammebetingelser. De valg som er tatt i utviklingen av datamaterialet, og forskernes beskjedne inngripen i utviklingen av dette i følge deres egne rapporter, må derfor kunne sees på som bevisste valg fra forskernes side.

Rapporteringene underveis i forsøksperioden har vist at forskerne har sett seg selv som en del av forsøksfasen ved å gi tydelige råd om forbedringer og endringer. De har derfor ansett evalueringens kontekst først og fremst å være innad i prosjektorganisasjonen. Valg av kontekst hos forskerne har, som det er vist, igjen følger for hvordan tilretteleggingen av læring skal skje. Ulike kontekster krever ulik tilrettelegging for læring, og tilsier endringer i språk, formidling, og organisering av evalueringen. Dette har følger for evalueringens plass både innenfor og utenfor organisasjonen. Med en annen forståelse av den kontekst forskerne stod i, ville øvrige valg i de didaktiske kategoriene kunne gitt evalueringen et helt annet utfall.

I følge Qvortrup innebærer tilrettelegging for læring at medlemmer av organisasjonen må ha rom for ulik type refleksjon ut i fra hvilke spørsmålsstillinger og typer evalueringer det handler om. Rammene gir rom eller begrensninger for denne refleksjonen, og har betydning for om evalueringen gis en rolle som refleksjonsverktøy. Hvordan evalueringen er tilrettelagt for dette har derfor igjen betydning for om den kan bidra til læring i en ”lærende organisasjon”.

## Arbeidsmåter

I evaluering av reform av trosopplæring er det, slik redegjørelsen viser, valgt en tydelig brukerstyrt evalueringsmetode som hører inn under substansmodeller i Vedungs taxonomi over ulike evalueringsmodeller. Som nevnt taler Vedung varmt for bruk av ulike type metode til ulikt bruk, og for å benytte større variasjon i metoder slik at det er mulig å dekke et bredere spekter av perspektiver i evalueringssammenheng. Jeg vil i det følgende beskrive hvilke valg forskerne i denne evalueringen har tatt, og hvilke andre valg de har hatt som kunne gitt andre perspektiver til evalueringsarbeidet. Det er lite hensiktsmessig for en enkelt evaluering, selv større evalueringer som denne, å benytte alle modellene som her presenteres. Det er heller ikke min intensjon ved å presentere alle. Likevel mener jeg det er grunn til å spørre om ikke forskerne hadde vært tjent med å benytte flere, eller i alle fall andre metoder enn de som er benyttet i denne evalueringen.

Evalueringsforskerne har valgt å gå utenom en måloppfyllelses modell, og har nok derfor valgt en modell som ligger nærmere en målløs resultatmodell. Men valget er også orientert om brukere, forventninger og behov, og hører derfor mest av alt inn under en brukerorientert evaluering. Fokus som disse modellene gir er tidligere omtalt i oppgaven som snever. Jeg vil i det følgende derfor kort vise hvilke andre muligheter som finnes, og hva de kunne gitt av andre perspektiver.

Helsysteminnrettet modell for hele organisasjonen ville egnet seg bra, fordi den i sterkere grad ville sørget for å belyse reformen ut fra hele organisasjonen Den norske Kirke, og ikke bare reformorganisasjonen. Deler av evalueringen er for øvrig innadrettet og hører mer hjemme under institusjonelle, profesjonelle modeller som har fokus på selv-evaluering, kollega bedømming og ekspert/sakkyndig vurdering. Den sist omtalte kunne med fordel vært benytte for å få mer fokus på hva faginstansene i organisasjonen kunne bidratt med av innspill. Så langt jeg kan se er evalueringen fri for økonomiske vurderinger, som egentlig også kunne vært interessante perspektiver i denne sammenhengen. Har det lønnet seg å iverksette forsøksfasen? Har den vært produktiv nok, i forhold til organisering og utfall/resultat? Burde man i større grad for eksempel ha satset på å hente inn faginstanser til å utvikle de nye tiltakene som skulle fungere lokalt, i stedet for at hver forsøks menighet skulle "finne opp kruttet" selv? Med de ressurser en mindre menighet har til rådighet vil det ikke være overraskende om en cost-benefit eller kostnads- effektivitets analyse av dette

perspektivet, ville medført at ressursene ville vært benyttet annerledes i en del av forsøksfasen.

Både interessentmodellen og framfor alt bieffektmodellen, ville kunne bidratt med helt andre perspektiver. Vedung argumenterer sterkt for valg av bieffektmodellen fordi den bidrar til å avdekke de uventede og uønskede effektene som ofte viser seg ved evalueringer. Ved hjelp av en slik modell ville andre sider ved reform av trosopplæring vært belyst. Valg av bieffektmodellen kunne gitt fokus på hva forsøkstiden har gjort med organisasjonen totalt sett. Når ressurstildeling til kirken har vært kanalisert gjennom dette forsøksprosjektet, hva har det da gått ut over? Hvilke områder av organisasjonen har eventuelt tapt på dette, og hva innebærer det nå og på sikt? Hva har evalueringen gitt av svar og føringer for grupper som blir involvert enten de vil eller ei? Hva har forsøkstiden bidratt med eller skapt utfordringer for når det gjelder bygging av organisasjonskultur? Hva har den betydd for avklaring av kompetanse som kreves, som igjen vil ha følger for profesjonene i organisasjonen? Dette breddeperspektivet savnes i den summative rapporten. Kanskje er det gjort med vilje for å sørge for en positiv vinkling? Bevilgningene til reformen var ikke endelig avklart, og det er grunn til å tro at forskerne var seg dette ansvaret bevisst.

Forskerne har brukt flere arbeidsmåter i sitt arbeid. Alt fra dokumentstudier, Innholdsanalyser av teologisk refleksjon, spørreskjemaundersøkelser og casestudier. De har foretatt systematiske gjennomganger av forsøksprosjektenes tiltaks- og årsrapporter, av reformens støttestrukturer, og gjort analyser av diverse statistikk, samt en spørreskjemaundersøkelse i 2007 til et utvalg av ikke-forsøksmenigheter. Forskerne har gjort grundig arbeid på et bredt grunnlag i forsøksfasen, og har lagt ned mye ressurser i å beskrive og oppsummere hva som har skjedd i de ulike fasene. Oversikten over studier viser at det er stor variasjon i arbeidet, og undersøkelsene favner relativt bredt i forholdt til å belyse hva som er skjedd i reformen. Likevel mener jeg at forskerne burde valgt innfallsvinkler og metoder til sin forskning som ville gitt større bredde i forskningen når det gjelder organisasjonen som helhet.

Hvis valget har vært å beskrive hva som skjer, og ikke aktivt være med å påvirke i evalueringsprosessen, så har forskerne slik de har lagt opp sin evaluering valgt mange gode løsninger for sitt arbeid. Hvis evaluering imidlertid også var ment å bidra aktivt til endring i organisasjonen, og fungere som mer enn en rapport, burde valg av arbeidsmåter og metode ha underbygget dette.

## **Gjennomføring**

Gjennomføring er lagt inn som en naturlig del av de didaktiske kategoriene i planleggingsfasen i modellen, fordi en tenkt gjennomføring av de øvrige definerte kategoriene er nødvendig å reflektere over som en del av planleggingen. Mye kan avdekkes i denne fasen som kan hindre problemer når evalueringens planer iverksettes for alvor. Hvorvidt forskerne i reform av trosopplæring har brukt tid på refleksjon rundt gjennomføringen på forhånd er vanskelig å si.

Gjennom arbeidet med denne kategorien kunne organisasjonen hatt en mulighet til å vurdere de bestemmelser som var gjort i forhold til de andre kategoriene i planleggingsfasen, og hvorvidt disse valgene fungerte godt og konsistent i forhold til hverandre. Det som er beskrevet om evaluering av reform av trosopplæring så langt i kapitlet gir grunn til å mistenke at det ikke er viet mye tid til denne fasen i forberedelsen.

## **Planlegging av evaluering**

Det er presentert en virkningsmodell for evalueringen i den første rapporteringen, samt framdriftsplan over evalueringens virkeperiode. Det foreligger også en oversikt over ulike arbeidsmåter som forskerne har tenkt å benytte. I så måte har forskerne lagt arbeid i planleggingsfasen for å organisere sitt arbeid.

Det er likevel tegn som tyder på at evalueringens planleggingsfase har vært viet for liten oppmerksomhet. Forberedelsene i evaluering av reform av trosopplæring preges slik jeg ser det av at ”veien er blitt til mens man er gått”. Dette kan være et bevisst valg av forskerne, og derfor en del av en utarbeidet plan for evalueringen i utgangspunktet. Rapportene viser at det er forsøkt å forbedre deler av evalueringens grunnlagsmateriale, som evalueringens resultater skulle måles ut i fra, underveis i prosessen. Dette materialet skulle besvare hvorvidt forsøksfasen har vært vellykket eller ikke, men forskerne har vurdert kvaliteten i materialet å være for dårlig til det.

Selv om dette materiale har vært for dårlig som grunnlagsmateriale er det ikke gjort forsøk på å hente inn nytt som erstatning for det. Det kan tyde på at det har vært for sent å gjøre forbedringer i materialet når evalueringen har vært i gang. Dette viser at det er nødvendig å

bruke god tid på kvalitetssikring av det verktøyet forskere skal bruke for å besvare mål og intensjon, samt tilrettelegge for refleksjon, endring og læring, i forkant av gjennomføringen.

Qvortrup har også vist at det å legge til rette for læring er en komplisert sak, som innebærer å ta hensyn til ulike nivåer for læring, der graden av refleksjon er viktig. Evne til refleksjon over egen praksis er nært knyttet til evne til læring. Derfor er planlegging og tilrettelegging av evalueringen vesentlig for å sørge for at det blir lagt til rette for refleksjon på ulike nivåer i organisasjonen.

Rom for refleksjon gjelder også forskerne selv. Evalueringsforskerne i reform av trosopplæring har gjennom sin følgeforskning rapportert underveis og offentliggjort sine rapporter på nett og holdt samlinger for presse og interesserte ved offentliggjøring. I så måte har de utsatt seg for kritikk underveis i arbeidet av andre forskere og fagmiljøer som har kjent til reformarbeidet og evalueringen av det. Likevel er mange av de samme argumentene som de selv har benyttet som betegnelse på svakheter ved reformen vært gjentatt underveis i hele forsøksfasen. Dette kan tyde på at de ikke har tatt tak i de utfordringene de selv har beskrevet som svakheter. Dette beskriver i en viss grad forskernes syn på egen rolle, og grad av refleksjon over eget arbeid. Det er også trolig at det sier noe om grad av ressurser som er nedlagt i planlegging og forberedelse.

Selve grunnlaget for læring gjennom evaluering legges i forberedelsesfasen av evalueringsprosessen. Hva er så skapt av grunnlag for læring i forberedelsesfasen av evaluering av reform av trosopplæring? Jeg mener at utfordringene som jeg har avdekket allerede fra startfasen av i forskernes arbeid innvirker på resten av forberedelsen, samt gjennomføring og resultatet av evalueringen. Dette vil jeg beskrive nærmere i gjennomgangen av selve gjennomføringsfasen av evaluering av reform av trosopplæring. Det er i selve gjennomføringen at konsekvensene av god eller dårlig planlegging vises.

Vedung legger vekt på at forberedelsesfasen bør avsluttes av en formell beslutning før gjennomføringsfasen iverksettes og evalueringen for alvor settes ut i livet. Som nevnt deler jeg dette synspunktet. Det er også trolig at forskerne som har evaluert reform av trosopplæring har fått godkjent sine planer for evaluering gjennom en formell beslutning av ledelsen i reformen, før evalueringen ble iverksatt. Forskerne kan i så fall ikke sies å være alene om ansvaret for de svakheter planleggingsfasen eventuelt burde ha avdekket.

### 5.1.2 Gjennomføringsfasen

Som omtalt i forrige kapittel så starter mange evalueringsoppdrag i gjennomføringsfasen. Slik jeg ser det er det sannsynlig at dette også i stor grad gjelder forskerne som har evaluert reform av trosopplæring. Rapporteringsgrunnlaget er produsert underveis, og er utviklet og bearbeidet gjentatte ganger i forsøksfasen. Dette mener jeg er uheldig både med tanke på evalueringens kvalitet og dens plass i organisasjonen. Dette har igjen betydning for hvilken posisjon evalueringen får i forhold til de målgrupper som er ment å lære av den.

Gjennomgangen av forberedelsesfasen over mener jeg viser at evalueringens rolle i organisasjonen vil bestemmes av hvilke forutsetninger den er gitt. Det som ikke er planlagt er det derfor vanskelig å ”hente inn” og tilrettelegge for i gjennomføringen. Den første fase i gjennomførelsesfasen innebærer å tilrettelegge for en rekke støtteprosesser som er nødvendige for å iverksette det som ble bestemt i forrige fase. Blant annet å sørge for at det finnes nok personalressurser til å utføre gjennomførelsesprosessen.

#### *For dårlig bemanning?*

Gjennom hele forsøksfasen har forskerne gitt reformledelsen tilbakemeldinger om behov for endringer vedrørende rapporteringsverktøyet. Samtidig har de ytret stor forståelse for sekretariatets pressede arbeidssituasjon, og som nevnt hevdet at det har vært behov for å styrke sekretariatet med flere stillinger. Jeg har tidligere hevdet at dette sier noe om forskernes rolle, og hvilket ansvar de har tatt for utvikling av sitt eget grunnlagsmateriale. Det er likevel Stortinget og de bevilgende myndigheter som her må ta det største ansvaret. Det faktum at det ikke forelå noen økonomisk ramme for hele forsøksfasen fra starten av i reformen, har gjort det vanskelig å planlegge. Reformledelsen visste ikke fra år til år hvor mye midler de hadde til rådighet, og har derfor hatt begrenset med muligheter til å legge føringer fram i tid for sin egen virksomhet. Det er likevel deres ansvar å tildele nok ressurser fra den totale ”potten” til kvalitetssikring og utføring av oppdraget. Forskerne har hatt ansvar for å informere reformledelsen om hva de økonomiske betingelsene deres har gitt av muligheter og begrensninger, mens ledelsen har tildelt dem de økonomiske rammene de har hatt. Reformledelsen må derfor ta ansvar for de betingelser forskerne er gitt, men de har ikke ansvar for at det ikke har vært mulig å legge bedre føringer fram i tid når det gjelder behov for økte ressurser fordi reformen vokste. Dette ansvaret må Stortinget bære.

Forskerne var ment å fungere som et verktøy for ledelsen i deres arbeide, og på den måten fikk de makt og muligheter til å påvirke forsøksarbeidet. Hvorvidt de har hatt denne funksjonen i praksis er et annet spørsmål. Funnene mine viser hvilke råd som er gitt og hvilke som er fulgt opp av prosjektledelsen gjennom forsøksfasen, og jeg har gjennom det stilt spørsmål ved hvilken rolle forskerne har hatt i reformorganisasjonen. Det er mye mulig at mine tolkninger i dette er vel strenge i forhold til den situasjonen reformledelsen - og derigjennom forskerne - har vært i når det gjelder de økonomiske rammebetingelsene. Jeg har belyst deres rolle og utførelse ut fra hvordan en ideelt sett bør utarbeide en evaluering for at den skal føre til læring. Jeg mener derfor at det er en berettiget kritikk av evalueringen, selv om ansvaret antakelig ligger nærmere Stortinget enn reformledelse og forskere.

En av mine foreløpige hypoteser i oppgaven var at kvaliteten i evalueringens utarbeidelse har betydning for hvilken rolle den senere får i organisasjonen. Men også synet på organisasjonen er utslagsgivende for hvilken rolle evalueringen får, fordi det har konsekvenser for evalueringens organisering og hvordan den bør forberedes og tilrettelegges. Hvordan reformledelse og medarbeidere ser på organisasjonen vil derfor være avgjørende for kvaliteten på gjennomføringen.

### *Kongresjonalistisk, episkopal, synodal organisasjon*

For å kunne bedømme hvorvidt en evaluering er gjennomført for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”, er det nødvendig å se på hva utførelsen sier om hvilken organisasjonsforståelse som preger organisasjonen.

Den norske kirke kan beskrives på svært mange måter. Ulike perspektiver eller begreper fanger opp ulike, karakteristiske sider eller fasetter ved kirkens identitet, organisering og virksomhet. Hvilket perspektiv som oppleves som aller mest betegnende for Den norske kirkes virksomhet og formål, vil være avhengig av ståsted og kirkesyn. Trossamfunnet Den norske kirke har flere kilder til sin identitet som fellesskap. Det er ikke uvanlig å si at det er tre ulike kirkeforståelser som konvergerer i måten Den norske kirke er organisert på. Det er vanlig å sonde mellom episkopale, synodale eller kongregasjonalistiske kirkeforfatninger. En episkopal kirke bygger på en hierarkisk og bispestyrt organisasjonsstruktur, hvor biskopens tilsynsembete (episkopatet) utgjør kirkens øverste organ. I en synodal kirke vil et kirkemøte (synoden), sammensatt av utsendinger fra menighetene, opptre som konstituerende organ for kirkesamfunnet. Både prester og lekfolk kan ha plass i synoden.

Etter kongregasjonalismens prinsipper vil den enkelte menighet være uavhengig og selvstyrt. Den norske kirke er, som mange andre kirker, forfatningsmessig en blandingsmodell: Den består av selvstendige menigheter, den har biskoper og den har et kirkemøte (Utdrag fra Staten og Den norske kirke, regjeringen.no).

I evaluering av reform av trosopplæring kan mye tyde på at et kongresjonalistisk syn på organisasjonen har vært dominerende i organisering og helhetlig tenkning hos både reformledelse og oppdragsgivere. Stortingets valg av ”bottom up”-strategi for innføring av reformen har etter mitt syn innvirket på alle valg av strategier, også evalueringsforskernes. Respekten for det lokale ledd – førstelinjetjenesten – i organisasjonen har vist seg i organisering og helhetlig tenkning gjennom hele forsøksfasen. De som initierte reformen, og ledelsen av forsøksfasen, har hatt sterk tro på at menigheten lokalt har hatt svaret på hva innhold i og organisering av trosopplæring i fremtiden skal være.

Forsøksfasen har båret preg av kreativt kaos, som i følge Nonaka og Takeuchi er et av fem punkter som er sentrale for kunnskapsutvikling i en organisasjon (Nonaka og Takeuchi 1995). Senge beskriver også i sin teori om utvikling av ”lærende organisasjoner” at organisasjonen må fungere som desentraliserte, anti-byråkratiske fleksible enheter, der alle medlemmene skal føle en forpliktelse og et ansvar for helheten i organisasjonen. Individuell læring er en grunnleggende forutsetning for kollektiv læring, samtidig som systemtenkning ved å skape forbindelser mellom personlig mestring, mentale modeller som inngrodde antakelser, felles visjoner, og gruppelæring gir arbeidet i den ”lærende organisasjonen” retning og mening (Senge 1999). I så måte har reformen vært utført i tråd med nyere organisasjonsteori som responderer med organisasjonsforståelsen ”lærende organisasjon”. Det er heller ikke her utfordringene i forsøksfasen ligger.

Organiseringen av forsøksfasen har gått utenom Kirken som organisasjon, og store deler av organisasjonen har derfor ikke deltatt i forsøket. Lokale prosjektledere og nyansatte har hatt den største innflytelsen, og øvrige tilsatte har kun vært involverte dersom de har hatt en rolle i forsøksprosjektet, eller de har hatt en prosjektledelse som har sett verdien av å drive helhetlig virksomhet lokalt. Forskningen viser at dette har vært personavhengig og kontekstuell, og at mange har slitt med nettopp integreringen av prosjektene i det øvrige arbeidet. Forskerne har omtalt forsøksarbeidet lokalt som ”en stat i staten” der det har vært konflikter, og jeg mener det kan være grunn til å hevde at hele reformen med sin



---

organisering og ledelse er forstått som en ”stat i staten” av tilsatte i kirkens øvrige organisasjon.

I følge Senge har aktørene og relasjonene/samhandlingen mellom dem innvirkning på organisasjonens evne til å lære. Disse kulturelle forutsetningene har derfor betydning for evalueringen og forsøksarbeidet generelt. Jeg mener det er grunn til å spørre om ikke hensynet til organisasjonens spesielle tredeling burde vært tatt mer på alvor, og faktisk kunne ha hatt positiv innvirkning på utvikling og gjennomføring av forsøksfasen i reformen. Jeg mener det er trolig at det ville ha betydd at fagmiljøer og tilsatte i organisasjonen ut over prosjektorganisasjonen vill tatt større del i reformarbeidet. Det ville igjen kunne hatt positiv innflytelse på innholdsbestemmelse av trosopplæringen, og det ville kunnet sørge for å ivareta bedre den utprøvende funksjonen som forsøksfasen burde hatt i større grad i følge forskerne. Det ville også sannsynligvis hatt stor betydning for implementeringen på sikt, fordi flere i organisasjonen ville følt ansvar og behov for å delta i reformen enn tilfellet er nå.

Selv om Den norske kirke er en spesiell organisasjon med tanke på den omtalte tredelingen i organisasjonsforståelse, er organisasjoner generelt avhengige av at implementering planlegges og sørges for i starten av en evalueringsprosess. Det handler om å legge til rette for størst mulig grad av involvering på ulike nivåer i hele organisasjonen. Vurdering og refleksjon må skje på flere arenaer og inkludere ulike aktører i organisasjonen. Endringene som er beskrevet i evaluering av reform av trosopplæring handler om mål, normer og verdier, blant annet i synet på barnet og innholdet i trosopplæringen, som må sies å handle om overordnede forhold i organisasjonen. Den læring man er ute etter å oppnå bærer derfor preg av å være dobbeltkretslæring (Argyris & Schön 1978). Denne formen for læring i en organisasjon forutsetter endringsvilje. Argyris & Schön vektlegger tilstrekkelig objektiv og kontrollerbar informasjon, frie valg basert på kunnskap, og felles forpliktelser over beslutninger for å sikre endringsvilje i organisasjonen.

Inkludering og involvering er en forutsetning for å oppnå endringsvilje, og det må organiseres slik at flest mulig opplever seg inkludert og involvert. Her kan det synes som om festtaler og begeistring har forsøkt å erstatte en reell tilretteleggelse for deltakelse. Involvering handler om tilrettelegging for deltakelse i alle deler av prosessen, slik den didaktiske relasjonsmodellen har vist. Denne involveringen er nødvendig forutsetning for en vellykket gjennomføring.

### *Gjennomført som virksomhet for læring?*

Den ”lærende organisasjon” ser på evaluering som instrumentell og opplysende, der evalueringens funksjon er å bidra til læring (Dahler-Larsen 1998). Jeg er usikker på om evalueringen her har hatt som mål å være en funksjon som bidrar til læring. Evalueringen er slik jeg ser det mer i tråd med evaluering tilhørende den institusjonelle organisasjon slik Dahler-Larsen beskriver den. I den institusjonelle organisasjon ses evalueringen på som en rite, der funksjonen er å revidere institusjonen. Evalueringen som rite signaliserer bestemte forhold i organisasjonen og utad. Ritualet meddeler grunnleggende forhold om verden; hva som eksisterer, hva som er riktig og forkjært, og hva som trengs å gjøres – alt sammen noe som kanaliserer oppmerksomhet og energi i bestemte retninger og ikke i andre.

Ritualiserende evalueringer kan derfor påvirke hva som oppfattes som etterstrebbare mål i en organisasjon; de kan styrke eller svekke bestemte former for solidaritet og samarbeid, og de kan gjøre bestemte typer av aktører tause eller talende. Det er dette som i følge Dahler-Larsen gjør evalueringer konstitutive og med på å konstruere virkelighet.

Evalueringen som rite har ingen dårlig funksjon i organisasjonen. Jeg er likevel uenig med Dahler-Larsen i at ritefunksjonen er den ideelle. Uten at evalueringer tar mål av seg til å føre til læring vil den ha langt mindre mulighet for å skape refleksjon og rom for læring hos aktørene i organisasjonen. Dette vil igjen medføre at den vil være mindre interessant i organisasjonsutvikling. I stedet for å legge føringer for åpenhet og kritisk refleksjon hos den enkelte medarbeider, gjøres her forskere til organisasjonsutviklere og besittere av sannheten. Hvis så blir tilfelle vil evaluering som instrument for læring kunne utsettes for det instrumentalistiske mistak (Skjervheim 1996), ved at agendaen med den er noe annet enn det den utgir seg for å være.

Ideelt sett skulle det materialet som er samlet inn i reform av trosopplæring kunne ført til både innholdsbestemmelse av trosopplæringen, og tilrettelegging for den systematisering som kreves for å oppnå dette, samt gitt et godt svar på om forsøksfasen har vært vellykket hvis dette var målet som evalueringen skulle vurdere ut i fra. Med bedre planlegging ville trolig valg av parametre også falt på noe annet. Dette forarbeidet ville kunne gitt evalueringen en lærende funksjon i organisasjonen. Bedre planlegging og mer prospektiv holdning til evalueringsarbeidet, ville også gitt gjennomføringen en sterkere posisjon i organisasjonen enn hva jeg vil anta er tilfelle nå.

Gjennomføringsfasen er også ment å sørge for refleksjon rundt den evalueringspraksis som er utført, som en del av den helhetlige evalueringsprosessen. Den didaktiske evalueringsmodellen legger opp til en revisjon av den praksiserfaring som er foregått i gjennomføringsfasen. Hvorvidt denne refleksjonen har skjedd, eller hvem som eventuelt har deltatt i den, er ukjent. Forskerne har som nevnt i liten grad omtalt sitt eget arbeid i rapporteringene.

### **5.1.3 Nykobling og metakoblingsprosess**

En av forutsetningene for at evalueringen utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”, er at evalueringen i seg selv utsettes for evaluering.

Nykoblingsfasen forutsetter det Dale vil kalle rasjonalitet i eget arbeid. Det vil si å utsette seg for kritikk, og ha en kritisk holdning til egen forskning som gjør det mulig å endre praksis i henhold til nye erfaringer. Det er her det største læringspotensialet ligger i evalueringer.

Deler av delrapportene og underveisrapportene i evaluering av reform av trosopplæring viser at det kan ha foregått flere nykoblinger. Der forskerne beskriver svakheter i materialet, kan dette tolkes som et forsøk på å se helhet og vurdere eget arbeide, samt legge til rette for forbedringer. Det er imidlertid ikke nedtegnet mye refleksjon direkte om forskernes rolle eller virksomhet i disse rapportene. Det betyr selvsagt ikke at det ikke har foregått. Tvert i mot er det trolig at erfarne forskere har vært opptatt av å utøve kritikk mot sitt eget og medforskeres arbeid.

Det er likevel lite trolig at forskerne og ledelsen av reformen i deres vurderinger av evalueringen, og i de eventuelle nykoblingsprosesser som har funnet sted, har hentet inn synspunkter fra en større bredde i organisasjonen. Hvis dette hadde vært tilfelle er det mer trolig at dette ville fått betydning for evalueringens utforming og organisering i den videre prosessen. Dette har ikke skjedd. Det er derfor vanskelig å svare på om evaluering av reform av trosopplæring har fungert etter hensikten. Den gir svar på bestillingen som ligger i avtalen med reformledelsen, men gir den svar på oppdragsgiver intensjon? Har den bidratt til å tilrettelegge for refleksjon og endring hos den enkelte aktør i organisasjonen?

I tillegg til denne nykoblingsprosessen anbefaler Vedung også å gjennomføre en prosess for metakontroll. Denne metakontrollen tar enda et skritt tilbake fra evalueringsobjektet og vurderer den nykoblingsprosessen som er utført, samt evalueringen i et videre perspektiv enn i det forrige. Jeg har koblet denne fasen i prosessen på Qvortrups 4.ordens vitensposisjon. Dette er som nevnt i følge Qvortrup ikke en individuell, men en kollektiv læringsprosess. Denne form for læring foregår når grupper av individer endrer forutsetninger for den læringen som foregår, eller når forutsetningene endres av noen utenfor. Evalueringens forutsetninger, forankring og karakter i møte med omgivelsene skal vurderes. De lærende i kirken, og deres opplevelse av hva evalueringen har utrettet for dem og fellesskapet/virkksomheten, skal være gjenstand for refleksjon i denne fasen. Evaluering bør som verktøy i en ”lærende organisasjon” ha som mål å bidra til å inkludere og involvere aktører/medlemmer i organisasjonen best mulig. Dette innebærer også å inkludere de kritiske røstene. Hvorvidt denne metakontrollen har skjedd, eller vil foregå i forbindelse med reform av trosopplæring, er ikke kjent.

På mange måter er den meta-evaluering jeg har foretatt i denne oppgaven et eksempel på et bidrag til meta-kontroll. Hvorvidt den vil fungere som kunnskapsproduksjon er avhengig av hvordan den blir tatt i mot og benyttet av forskerne i evaluering av trosopplæring, reformledelsen, og i organisasjonen som helhet.

### *Oppsummering*

I dette kapitlet har alle fasene i en evaluering vært belyst gjennom den didaktiske evalueringsmodellen som ble presentert i forrige kapittel. Modellen er brukt som utgangspunkt i analysen av evaluering av reform av trosopplæring, og selve drøftingen av hvorvidt evalueringen er utarbeidet og gjennomført for å gi et best mulig grunnlag for læring i en ”lærende organisasjon”. I arbeidet med denne oppgaven har jeg forsøkt å holde fokus både på den normative og den deskriptive siden ved evalueringer<sup>4</sup>. Jeg har både evaluert forskningen i reform av trosopplæring og søkt i teori å finne svar på hvordan en evaluering ideelt sett bør utarbeides og gjennomføres for å føre til læring. Dette har jeg så gjort ved

---

<sup>4</sup> I et pedagogisk studium vil vi dels kunne forsøke å gi en beskrivelse av den didaktiske relevante viten vi har og av verdier og normer som er eller har vært sentrale i undervisnings- og eller oppdragelsessammenheng. Angriper vi på den annen side didaktikken ut i fra spørsmålet om hvordan vi bør handle i en gitt situasjon, er vi derimot på jakt etter det foreskrivende element i didaktikken. Også i dette tilfelle må vi både ha viten og verdier som bakgrunn for de konklusjonene vi trekker (PFI 1975, s2). ”På den måten kan vi selv bli mer bevisst på hvilke grunnlag vi har for våre didaktiske overveielser og avgjørelser” (PFI 1975, s 3). I det virkelige liv går disse aspektene ofte over i hverandre (Engelsen 1990).

hjelp av en didaktisk tilnærming til evalueringsfeltet. I følge Imsen bør ikke normativ didaktikk og analytisk didaktikk oppfattes som motsetninger. De gir hver på sin måte bidrag til å forstå ulike sider ved lærerens (les evalueringsforskeres) arbeid (Imsen 1999, s 42).

Den didaktiske evalueringsmodellen viser at en didaktisk innfallsvinkel til evalueringer gir fokus på planlegging og forberedelse, som igjen kan sikre kvalitet i utarbeidelse, gjennomføring og resultat av evalueringer. Modellens mange faser viser også hvordan evalueringen bør tilrettelegges for å sikre at den skal føre til læring for den enkelte aktør og derigjennom hele ”den lærende organisasjonen”.

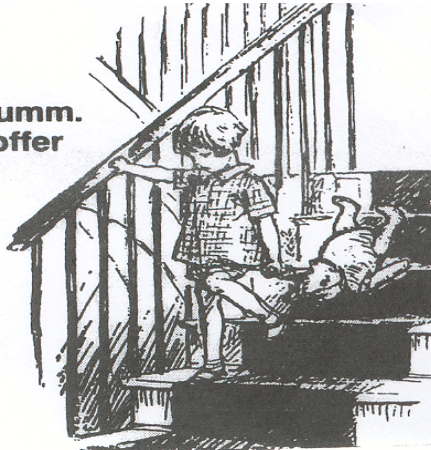
Gjennomgangen av den didaktiske evalueringsmodellen og de ulike fasene i evalueringsprosessen, og den måten evalueringen er løst på, har vist at veldig mye av kvaliteten i evalueringen som helhet avgjøres i forberedelsesfasen. Det som ikke er tatt hensyn til eller tilrettelagt for i denne delen av evalueringsprosessen kan ikke forventes å få innflytelse senere i prosessen.

Analysen har avdekket flere svakheter ved tilretteleggingen for læring i evaluering av reform av trosopplæring, og jeg mener det i hovedsak handler om hvordan evalueringen er organisert og hvilken rolle den og forskerne har hatt i organisasjonen. Det refererer derfor til evalueringens rolle som kunnskapsproduksjon, som igjen henger sammen med evalueringens kvalitet. Ved bedre tilrettelegging fra starten av mener jeg at evalueringen kunne hatt en viktigere funksjon som bidragsyter til å nå de mål reformledelse og Storting hadde med reformen. Det ville også gitt bedre kvalitet i det innhentede datamaterialet som skulle danne grunnlaget for den summative evalueringen. Den ville høyst sannsynlig også hatt en bredere målgruppe og en innfallsvinkel til hele organisasjonen som ville gitt bedre implementering. Bedre kvalitet i produkt, ville gitt evalueringen en sterkere posisjon i organisasjonen, fordi den da hadde levert svar på reformens mål og utfordringer, og sørget for involvering og inkludering av organisasjonen som helhet.

Jeg mener derfor at testevalueringen min underbygger mitt anliggende om å vektlegge en mer prospektiv innfallsvinkel til evalueringsfeltet. Det er vanskelig å få læring ut av en evaluering som ikke er tilrettelagt for det i organisasjonen, og det er vanskelig å få læring til å bli en ønsket endringsprosess hvis de ønskede parametre ikke er definerte i forberedelsesfasen, slik de didaktiske kategoriene bidrar til.

## 6. Oppsummering/konklusjon

– Hvor går vi, spurte Ole Brumm.  
– Ingen steder, svarte Kristoffer Robin.  
Og så gikk de dit.



Gjennom arbeidet med analysen av evaluering av reform av trosopplæring har følgende overskrift tredd fram: ”veien blir til mens vi går”. Evalueringsforskerne har evaluert et kreativt kaos, der forsøket på å bringe orden, samt skape resultater ved hjelp av systematikk, har vært målet. Forsøkets organisering og gjennomføring har på mange måter vist sterk tro på den kunnskapen som er utviklet gjennom lange tradisjoner i organisasjonen. Troen på at de lokale leddene i organisasjonen ville finne fram til den videre vei i trosopplæringen har vært stor. Menighetene ble sett på som det naturlige utgangspunkt og basis for utvikling og nyskapning. Utfordringene for reformledelsen er at svaret på innholdsbestemmelse, og hva Den norske Kirkes trosopplæring skal være i framtiden, ikke er gitt gjennom forsøks tiden. Trosopplæringens kontekstualitet ble stadfestet, men var vel egentlig en kjent kunnskap fra før, og heller grunnen til at behovet for reformen meldte seg i første omgang. Den kongresjonalistiske linjen i organisasjonen har vært dominerende, og tilsynelatende stått i veien for faginstansenes påvirkning og innflytelse. Sterkere posisjon fra faginstanser ville gjort at forsøksfasen ville gitt tydeligere resultater, både i forhold til faglig tyngde, implementering og organisasjonsutvikling. Dette har forskerne sett, og forsøkt å vinne fram med synspunkter for, men i liten grad blitt hørt i.

Følgforskningen har på mange måter vært følge med på veien i forsøksarbeidet, og ”reiseskildringene” til forskerne har vært gode rapporter over det som er skjedd. Likevel har evalueringen bare i mindre grad bidratt til endringer og kunnskapsproduksjon i reformen. Dette viser funnene av hva forskerne har fått gjennomslag for, og hvilken kvalitet det har vært på det materialet som forskerne har benyttet som grunnlag for sine vurderinger.

Mine egne erfaringer med evalueringer har vist meg at det er en lite uttalt, men likevel tydelig forutsetning for evalueringer at de må kunne vise vei i videre utvikling. Men det er omstridt hva evalueringens veiviserfunksjon skal bestå i. På den ene siden har jeg hevdet at det er dumt hvis evalueringer med dårlig kvalitet blir brukt og får effekt i organisasjonen, og på den annen side er evalueringer som ikke skaper endringer helt meningsløse. Er det ikke en motsetning i dette?

Det er i dette spenningsfeltet utviklingen av evalueringsfeltet befinner seg. På et eller annet punkt peker evalueringsforskere ut retningen for evalueringen. Hva som skaper disse retningsvalgene, og hvordan den standard og verdi evalueringen skal måle ut i fra bestemmes, er en oppgave for forskere. Evalueringsforskning blir derfor et spørsmål om å finne fram til de gode verdikriteriene som innebærer at de veivalg som tas skaper validitet og kunnskapsproduksjon.

Skal evalueringen ha som mål å føre til læring, eller skal den kun være en beretning eller rapport om hva som er skjedd? Hvor går skillet mellom en god og en dårlig evaluering? Er det noen felles kriterier å forholde seg til? Deskriptive beskrivelser alene er å regne som rapporter over en historisk hendelse. Den bør være en redegjørelse, som ikke spør om hvorfor det som er skjedd har skjedd. For meg går skillet mellom rapport og evaluering akkurat her. En evaluering spør hvorfor, og er opptatt av å beskrive dette. En evaluering er derfor også normativ. Hvis det skal være noen grunn til å spørre hvorfor, så innebærer evalueringens rolle også en intensjon om at det skal føre til refleksjon og læring i en eller annen form. Dette krever igjen tilrettelegging. Det normative aspektet stiller derfor store krav til forskernes utførelse av evalueringen i utarbeidelse og gjennomføring.

Vedung har vist at evaluering innebærer i etterkant å oppdage, kartlegge, og bedømme resultater med det formål å oppnå økt selvrefleksjon, dypere forståelse og gjennom dette ta bedre begrunnede beslutninger (Vedung 1998). For å oppnå økt selvrefleksjon og dypere forståelse forutsetter det at det har foregått prosesser av læring. Qvortrup viser med sine

vitensposisjoner at refleksjon forutsetter prosesser av læring, som krever ulike typer tilrettelegging i organisasjonen dersom refleksjon og læring skal finne sted.

For at evaluering skal innebære refleksjon over praksis, og føre til dypere forståelse, må evalueringen tilrettelegges for at læring skal finne sted. Det ligger også som en forutsetning i dette, at uten denne tilretteleggingen vil ikke evalueringer fungere etter sin intensjon i henhold til Vedungs forståelse av evaluering. Det er derfor nærliggende å hevde at med bakgrunn i denne forståelsen av evaluering og læring, er det et verdikriterium i all evaluering at de bør ha som mål å føre til læring.

Tilrettelegging for læring er et spørsmål om didaktikk. Det er denne tilretteleggingen for læring i organisasjonen jeg mener den didaktiske evalueringsmodellen kan fungere som et nyttig verktøy for. For at evalueringer skal fungere etter sin intensjon og bringe aktører og organisasjoner til dypere forståelse og refleksjon over praksis, må den legge til rette for refleksjon på ulike nivåer i "en lærende organisasjon". Dette er nødvendig for å kunne gi retning for veien videre.

Problemstillingen i oppgaven har vært: **Hvordan bør en evaluering utarbeides og gjennomføres for å gi et best mulig grunnlag for læring i "en lærende organisasjon"?**

Jeg har forsøkt å besvare den ved hjelp av anerkjent teori om evaluering og læring, og ved å utføre en empirisk undersøkelse i en meta-evaluering over en nylig gjennomført evaluering i offentlig sektor. Meta-evalueringen har vært både deskriptiv og normativ. Testevalueringen kan bare i svært begrenset grad benyttes som belegg for mitt anliggende, siden den ytre validiteten er så svak. Likevel har jeg tillatt meg i analysen å hevde at sterkere fokus på planlegging og tilrettelegging for læring på alle nivåer i organisasjonen, ville betydd at evalueringen i reform av trosopplæring sannsynligvis kunne bidratt til bedre resultat for forsøksfasen, og et bedre utgangspunkt for implementeringen i organisasjonen. Jeg har gjennom analysen og gjennomgangen av teorien vist til at en prospektiv innfallsvinkel til evalueringer kan bidra til å sikre kvalitet og validitet i resultater, og gjennom dette sikre evaluerings rolle som kunnskapsproduksjon.

Det å være på reise uten kart eller retning for reisen, bringer deg ikke uten videre dit du vil. Mange forutsetninger skal være innfridd for at evalueringer skal gi retning for reisen, og for at erfaringene skal være noe å ta med seg videre og bli en del av den nye kunnskapen som



tilegnes. Hva gjør at noen evalueringer får betydning og bidrar til endringer i organisasjonen, og andre ikke?

Godt didaktisk arbeid ved hjelp av god planlegging sikrer kvalitet i utarbeidelse og gjennomføring, og kan gi evalueringen funksjon som verktøy for læring i ”en lærende organisasjon”. Gjennomgangen av den didaktiske evalueringsmodellen har vist at det skal omfattende tilrettelegging i organisasjonen til dersom evalueringer skal føre til læring. Denne tilretteleggingen av evalueringen er med på å bestemme retningen for evalueringen, hva den skal måle, hva det skal måles med, for og med hvem, og hvordan målingen skal foregå. Arbeidet med de didaktiske kategoriene bidrar derfor til å bestemme de øvrige verdikriteriene i evalueringen.

I tillegg til dette bidrar den didaktiske evalueringsmodellen til å kvalitetssikre evalueringen i sin helhet, ved å vise til de ulike fasene som en evalueringsprosess innebærer. Med modellen som utgangspunkt for utarbeidelse av et program for evaluering, og ved å følge programmet og identifisere og utarbeide verktøy for å ivareta intensjonene for de didaktiske kategoriene, vil evalueringen bedre sikres god kvalitet og rolle i organisasjonen. Den vil da også gis gode forutsetninger for å fungere som kunnskapsproduksjon.

En av de fem forutsetningene til Nonaka og Takeuchi for kunnskapsoverføring er at kunnskapen som skal tilegnes må være mest mulig felles hos medlemmene i organisasjonen. Dette beskrives som særlig viktig i forhold til å forstå veien fra taus kunnskap til eksplisitt kunnskap, som igjen blir regnet som grunnleggende for å forstå kunnskapsoverføring. Kunnskapsoverføringen er i følge dem læring i prosess, der ulike nivåer av kunnskap overføres i trinn fra individ til grupper/team, som igjen fører til organisasjonskunnskap (Nonaka og Takeuchi 1995). Valg av arbeidsmåter og metode bør understøtte behovet for tilrettelegging i en ”lærende organisasjon”. Qvortrup har vist at ulike former for læring krever ulik tilretteleggingen overfor aktørene og i organisasjonen som helhet.

I organisasjonslæring er det naturlig å se på en ”lærende organisasjon” som organiserte læringssystem, der dens mål og mening er å organisere læreprosesser. Når fokuset er å systematisere arbeidserfaringer i organisasjonen som helhet, får vi erfaringslæring. Når fokuset er spesielt tilrettelagte læreprosesser, får vi læringserfaringer (Dale 1998).

Inkludering og deltakelse er en forutsetning for at læring skal finne sted. ”Fortell meg og jeg vil glemme, vis meg og jeg vil huske, involver meg og jeg vil forstå” (Confucius rundt 450

f.Kr ). Spørsmålet for evalueringsforskere blir da først og fremst hvordan evalueringen kan besørge best mulig inkludering for alle aktørene i organisasjonen, for på den måten å sørge for at læring kan finne sted.

Evnen til rasjonell fornyelse avhenger av medlemmenes kritiske analyser av deres rasjonelle atferd. Rasjonalitet ytrer seg i atferdsformer som det er gode grunner for. Rasjonelle handlinger og ytringer er forbundet med krav om gyldighet. Organisasjonsdidaktikken må derfor i følge Dale knyttes til språk- og kommunikasjonsteori som åpner for kritiske analyser. Kommunikasjon som kritiske analyser blir det overordnede grunnlaget for rasjonell læring og fornyelse (Dale 1998). Dale beskriver også evaluering som en viktig forutsetning for læring i "en lærende organisasjon" ved følgende formulering: "Om undersøkelser blir et kvalitetstrekk ved praksisen, vil det øke sjansene for at yrket kan bli en kilde til selvfornyelse" (Dale 1998).

Didaktikken er av flere teoretikere betegnet som opposisjonsvitenskap. Engelsen bruker heller betegnelsen "refleksjons- og problematiserings vitenskap" (Engelsen 1993). Dette fordi det har fått et faglig innhold – og ikke bare et utdanningspolitisk innhold. Dette gjør det lettere for didaktikken å være refleksjons-, problematiserings-, eller opposisjonsvitenskap – ikke bare støttevitenskap eller tjenesteytende vitenskap (Dale 1993). Jeg håper ved denne oppgaven å ha bidratt til å videreføre denne tanken når det gjelder evalueringer – ved å gi didaktikken en rolle også i evalueringsfeltet. En av didaktikkens fremste oppgaver i følge Engelsen, er å være en problematiseringsvitenskap som, ut i fra sin forskningsbaserte kunnskap kan analysere og problematisere de forhold som angår "undervisning" eller opplæringsprogram (Engelsen 1993). Jeg mener med min oppgave å ha argumentert for å utvide dette perspektivet til å omhandle evalueringsfeltet som opplæringsarena.

Siden kriteriespørsmålet er et lite avklart område i evalueringsforskning, er det etter mitt syn behov for å drive forskning for å utrede dette nærmere. En naturlig oppfølging av min problemstilling ville være å foreta en metasyntese over flere ulike evalueringer for å se om jeg ved hjelp av den didaktiske evalueringsmodellen ville lært noe mer om læring som kriterium, og hvorvidt de didaktiske kategoriene utfyller læring som kriterium, slik at den didaktiske evalueringsmodellen faktisk kan bidra til å løse verdikriteriespørsmålet i evaluering.

Det kunne også vært interessant å vurdere de funnene jeg har gjort opp mot andre sider ved evalueringsfeltet. Er for eksempel læringskriteriet like viktig i alle typer organisasjoner og organisasjonsforståelser? Er det en sammenheng mellom organisasjonsforståelse og læringsaspektet i evalueringer? Blir evalueringer benyttet som feedback i organisasjoner, og i så fall: hvilke typer feedback representerer de? Her er det mange forskningsoppgaver som det kunne vært spennende å ta tak i.

Jeg har selv vært på reise gjennom denne oppgaven, både i evalueringsfeltet og i oppgaveskriving. Oppgaveskrivingen i seg selv har vært en øvelse, der jeg har trøstet meg mye underveis med Steinsholts pedagogiske paradoks om læring:

*En student kan ikke med en gang forstå hva hun har behov for å lære. Hun kan bare lære det ved å utdanne seg selv, og kan utdanne seg selv bare ved å forsøke å gjennomføre det hun ennå ikke forstår (Kjetil Steinholt 1989:66).*

## Kildeliste

Alvesson, Mats og Sköldbberg, Kaj (1994): Tolkning og refleksjon

Argyris, C., & Schön, D. (1978): Organizational learning: A theory of action perspective

Blichfeldt, Jon Frode og Aamodt Per O.(1999): Om følgeforskning i Reform 94. I:Kvalsund, Rune m.fl (red.): Videregående opplæring - ved en skilleveg? Oslo:Tano.

Bjørndal, Bjarne og Lieberg, Sigmund (1978): Nye veier i didaktikken.

Dierkes, Meinolf, Antal, Child & Nonaka (2001): Handbook of Organizational Learning and knowledge. OUP.

Dale, Erling Lars (1992): Pedagogisk filosofi

Dale, Erling Lars (1998): Utdanning med pedagogisk profesjonalitet.

Dale, Erling Lars (1989): Pedagogisk profesjonalitet : om pedagogikkens identitet og anvendelse

Dale, Erling Lars (2001): Om utdanning

Dale, Erling Lars (2005): Kunnskapsregimer i pedagogikk og utdanningsvitenskap

Dale, Erling Lars (2006): Oppdragelse i det refleksivt moderne

Dale, Erling Lars (2008): Fellesskolen: Skolefaglig læring for alle

Dale, Erling Lars (2009): Læreplan

Dahler-Larsen, Peter: Den rituelle refleksjon – om evaluering i organisationer. Odense Universitetsforlag, 1998.

Dysthe, Olga (2001): Dialog, samspel og læring. Abstrakt forlag AS

Eisner, Elliot W. (1994): The educational Imagination. On the Design and Evaluation of School Programs

- 
- Gadamer, Hans-Georg: Truth and Method. 2nd rev. edition. Trans. J. Weinsheimer and D.G.Marshall. New York: Crossroad, 1989. ISBN 978-0826476975 excerpt
- Haug, Peder & Schwandt, T.A. (red.) (2003): Evaluating Educational Reforms
- Haug, Peder (2004): Sentrale resultat frå evalueringa av reform 97. I: Norsk pedagogisk tidsskrift, nr 2
- Kvernbekk, Tone (2005): Pedagogisk teoridannelse. Bergen: Fagbokforlaget
- Lund, Thorleif (2002): Innføring i forskningsmetodologi
- Monsen, Lars NPT 2004 nr 4
- Nespor, Jan (2002): Networks and Contexts of Reform
- Nordhaug, Odd (2002): Kunnskapsledelse
- Nowotny, Helga Scott, Peter Gibbons, Michael (2007): Rethinking Science: Knowledge and the Public
- Qvortrup, Lars (2004): Det vidende samfund - mysteriet om viden, læring og dannelse. Unge pædagogers serie B 73.
- Säljö, Roger (2005): Lärande & kulturella redskap – Om lärsprocesser och det kollektiva minnet. Norstedts Akademiska Förlag.
- Schriener, Michael (2003) :*Evaluation Theory and Metatheory*. I Kellaghan, T. & Stufflebeam, D.L. (Eds.): International Handbook of Educational Evaluation. Dordrecht: Kluwer, s. 15 – 30.
- Stake, Robert (2000): Advances in Program Evaluation
- Tyler, W. Ralph (1949): Basic Principles of Curriculum and Instruction
- Ulstrup Engelsen, Britt (2006): Kan læring planlegges?
- Vedung, Evert (1998): Utvärdering i politik og förvaltning. Lund.

Pedersen, Lene Tanggaard (2004): "Evaluering som regulering af læring", Artikel i 'Tidsskrift for arbejdsliv' nr. 4 2004, side 57-75.

Nettsteder:

[http://www.uhr.no/documents/Norsk\\_utdanningsforskning\\_Notat\\_til\\_KD\\_Januar\\_2007.pdf](http://www.uhr.no/documents/Norsk_utdanningsforskning_Notat_til_KD_Januar_2007.pdf)

[http://www.nifab.no/artikler/hva\\_er\\_forskning](http://www.nifab.no/artikler/hva_er_forskning)

<http://www.norad.no/>

[http://www.utdanningsdirektoratet.no/Artikler/\\_Lareplaner/Den-generelle-delen-av-lareplanen/](http://www.utdanningsdirektoratet.no/Artikler/_Lareplaner/Den-generelle-delen-av-lareplanen/)